

1299 PENNSYLVANIA AVE., NW
WASHINGTON, DC 20004-2402
PHONE 202.783.0800
FAX 202.383.6610

A LIMITED LIABILITY PARTNERSHIP



October 10, 2000

Commissioner for Patents Washington, DC 20231

Box Patent Application

Re:

U.S. Non-Provisional Utility Patent Application

Application No.: To Be Assigned

Filed:

Herewith

Title:

Annotated Plant Genes

Inventor:

David K. KOVALIC et al.

Atty. Dkt. No.: 04983.0204.CPUS00/38-21(15097)D

Sir:

The following documents are forwarded herewith for appropriate action by the U.S. Patent and Trademark Office:

- 1. Utility Patent Application Transmittal (PTO/SB/05);
- 2. U.S. Utility Patent Application entitled:

Annotated Plant Genes

and naming as inventors:

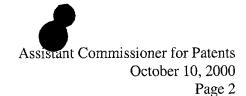
David K. KOVALIC and Jingdong LIU,

the application consisting of:

- a. A specification containing:
 - (i) <u>59,101</u> pages of a description prior to the claims;
 - (ii) 2 pages of claims (10 claims); and
 - (iii) a one (1) page abstract;
- 3. A CD-ROM containing the sequence listing;
- 4. Petition to Suspend Sequence Rules (in duplicate);







- 5. Our Check No. 337505 in the amount of \$130.00 to cover the petition fee; and
- 6. Two (2) return postcards.

This application is being filed without an executed Declaration, and without payment of official fees.

To facilitate the processing of the accompanying application, the CD-ROM containing the sequence listing can be located in Box 1. In addition, again to facilitate processing of the accompanying application, the pages setting forth the claims can be found in Box 17.

The U.S. Patent and Trademark Office is hereby authorized to charge any fee deficiency, or credit any overpayment, to our Deposit Account No. 08-3038. A duplicate copy of this letter is enclosed.

It is respectfully requested that, of the two attached postcards, one be stamped with the filing date of these documents and returned to our courier, and the other, prepaid postcard, be stamped with the filing date and unofficial application number and returned as soon as possible.

Sincerely,

David R. Marsh (Reg. No. 41,408) June E. Cohan (Reg. No. 43,741)

Enclosures

Please type a plus sign (+) inside this box

PTO/SB/05 (2/98) oved for use through 09/30/00. OMB 0651-0032 coffice: U.S. DEPARTMENT OF COMMERCE

Patent and Tra Under the Paperwork Reduction Act of 1995, ersons are required to respond to a collection of information as ss it displays a valid OMB control number.

UTILITY		
PATENT APPLICATION		
TRANSMITTAL		

Attorne	y Docket No.	04983.0204.CPUS01/3	38-21(15097)D	
First Named Inventor or Application Identifier			KOVALIC	
Title	Annotated Plant Genes			
Evenyage	Mail Lobal No			

PATENT APPLIC		First Named Inventor or Application Identifier KOVALIC		
TRANSMITTAL (Only for new nonprovisional applications under 37 CFR 1 53(b))		Title Annotated Plant Genes		
		Express Mail Label No.		
APPLICATION EI See MPEP chapter 600 concerning utili		Assistant Commissioner for Patents ADDRESS TO: Box Patent Application Washington, DC 20231		
1. *Fee Transmittal Form (For (Submit an original and a duplic		6. Microfiche Computer Program (Appendix)		
Specification (preferred arrangement set forth	[Total Pages below) ention ed Applications	7. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary) a. Computer Readable Copy		
- Statement Regarding Fed - Reference to Microfiche A	appendix	b. Paper Copy (identical to computer copy)		
- Background of the Inventi - Brief Summary of the Inve		c. Statement verifying identity of above copies		
- Brief Description of the D	rawings (if filed)	ACCOMPANYING APPLICATION PARTS		
- Detailed Description		8. Assignment Papers (cover sheet & document(s))		
- Claims - Abstract of the Disclosure		9. 37 CFR 3.73(b) Statement Power of (when there is an assignee) Attorney		
3. Drawing(s) (35 USC 113)	[Total Sheets] 10. English Translation Document (if applicable)		
4. Oath or Declaration	[Total Pages] Information Disclosure Copies of IDS Statement (IDS)/PTO-1449 Citations		
a. Newly executed (original control of the control	nal or copy)	12. Preliminary Amendment		
(for continuation/divis	plication (37 CFR 1.63(d)) ional with Box 17 completed)	13. Return Receipt Postcard (MPEP 503) (Two) (should be specifically itemized)		
i. DELETION OF I Signed statement atta	Box 5 below] NVENTOR(S) ched deleting inventor(s) named on, see 37 CFR 1 63(d)(2) and	14. *Small Entity Statement filed in prior application, Status still proper and desired		
A COLUMN		15. Certified Copy of Priority Document(s) (if foreign priority is claimed)		
Incorporation By Reference (useable if Box 4b is checked. The entire disclosure of the prior application, from which of the oath or declaration is supplied under Box 4b, is co as being part of the disclosure of the accompanying appliand is hereby incorporated by reference therein.		h a copy Check #337505 in the amount of \$130.00		
17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:				
Continuation Divisional Continuation-in-part (CIP) of prior application No: 09/654,617 (unofficial serial number)				
Prior	Application Information: Ex	xaminer: N/A Group/Art Unit: N/A		
18. CORRESPONDENCE ADDRESS				
Customer Number or Bar Code Label 22930 or Correspondence address below (Insert Customer No. or Attach bar code label here)				
David R. Marsh				
NAME HOWREY SIMON ARNOLD & WHITE, LLP				
Box No. 34				
ADDRESS 1299 Pennsylvania Avenue, N.W.				
CITY Washington	STATE	DC ZIP CODE 20004-2402		
COUNTRY US	TELEPHONE	202-783-0800 FAX 202-383-7195		
Name (Print/Type) David R. Marsh	June E. Cohan	Registration No (Attorney/Agent) 41,408/43,741		
gnature Date October 10, 2000 when Hour Statement: The forests actions and take 0.2 hours to complete. Time will vary depending months needs of the individual case. Any comments on the amount of time will are required to complete this				

30

5

10



ANNOTATED PLANT GENES

CROSS REFERENCE TO RELATED APPLICATION

The present application is a continuation-in-part of Application No: 09/654,617 (unofficial serial number) (Attorney Docket No: 04983.0204.00US00/38-21(15097)D), entitled "Annotated Plant Genes," filed September 5, 2000, inventors David K. Kovalic and Jingdong Liu.

FIELD OF THE INVENTION

The present invention is in the field of plant biochemistry. More specifically the invention relates to nucleic acid sequences from plant cells, in particular, nucleic acid sequences from maize, teosinte, soybean, *Arabidopsis*, cotton, sorghum, rice and wheat. The invention encompasses nucleic acid molecules that encode proteins and fragments of proteins. In addition, the invention also encompasses proteins and fragments of proteins so encoded and antibodies capable of binding these proteins or fragments. The invention also relates to methods of using the nucleic acid molecules, proteins and fragments of proteins, and antibodies, for example for genome mapping, gene identification and analysis, plant breeding, preparation of constructs for use in plant gene expression, and transgenic plants.

BACKGROUND OF THE INVENTION

The identification and isolation of plant genes belonging to biochemical and regulatory pathways are important in the development of nutritionally and agriculturally enhanced crops and products. Such nucleic acid molecules can be used in a variety of applications. For example, a nucleic acid molecule or a collection of nucleic acid molecules may act as a marker for a developmental or commercially valuable trait such as disease resistance. Additionally, they may be used to obtain homologues in the same or a different species. Nucleic acid molecules comprising coding sequences may also aid in gene expression studies that allow the dissection and elucidation of commercially useful traits.

The present invention provides nucleic acid molecules that are drawn from maize, soybean, rice, cotton, sorghum, wheat *Arabidopsis* and teosinte. They exhibit significant homology with known nucleic acid sequences belonging to a variety of biochemical and regulatory pathways.

Descriptions of biochemical and regulatory pathways are available from a number of well known academic and research organizations. An exemplary listing of such pathways can be

30

5

10



found in US patent application Serial number 09/371,146, the entirety of which is herein incorporated herein by reference, see especially pages 1-312. Additionally, several web sites and databases contain information pertaining to biochemical pathways and regulatory pathways. Examples of such web sites or data bases include: http://cgsc.biology.yale.edu (the CGSC maintains a database of E. coli genetic information, including genotypes and reference information for the strains in the CGSC collection, gene names, properties, and linkage map, gene product information, and information on specific mutations); http://www.labmed.umn.edu (the University of Minnesota's Biocatalysis/Biodegradation web page provides a search engine for compounds, enzymes, microorganisms, chemical formulas CAS registry, EC accession and microbial biocatalytic reactions and biodegradation pathways primarily for xenobiotic, chemical compounds such methionine, and threonine); http://wit.mcs.anl.gov/WIT2 (this website provides a functional overview which outlines metabolic pathways for organisms such as E. coli); http://ecocyc.PangeaSystems.com/ecocyc/ecocyc.html (this web site provides an overview of an E. coli metabolic map); http://www.biology.UCSD.edu (this web site provides information on signal transduction in higher plants); http://geo.nihs.go.jp (the Japanese National Institute of Health Science server provides information particularly on cell signaling networks); http://gifts.univ-mrs.fr (the Gene Intereactions in Fly Trans-world Server provides information on gene interactions, mostly centered on Drosophila gene interactions); http://sdb.bio.purdue.edu (this web site provides a data base of Drosophila genes); http://genome-www.stanford.edu (Stanford Genomic Research web site provides information on for example, Sacchromyces and Arabidopsis); http://www.psynix.co.uk (this web site provides illustrations and computer models of various cytokinins); http://www.sdsc.edu/Kinases/pk_home.html (this web site provides information on the protein kinase family of enzymes); http://transfac.gbf-braunschweig.de (the GBF web site provides information on regulatory genomic signals and regions, in particular those that govern transcriptional control); http://www.gcrdb.uthscsa.edu (this web site provides information on G-protein coupled receptors); http://www.biochem.purdue.edu (this web site provides information on secondary metabolism in Arabidopsis); http://home.wxs.nl/~pvsanten/ mmp/mmp.html (this web site provides a flow chart of metabolic pathways); http://www.genome.ad.jp/kegg/ regulation.html (this web site, the KEGG regulatory pathways web site, provides pathway maps, ortholog group tables, and molecular catalogs searchable data bases by enzyme, pathway, or EC

30

5

10



number); http://capsulapedia.uchicago.edu/ Capsulapedia/Metabolism/RegExpMet.shtml (this website provides expression information); http://www.zmbh.uni-heidelberg.de/M_pneumoniae/genome/META/ALL_META.GIF (this web site provides a graphic of metabolic pathways and the ways these pathways interact); http://moulon.inra.fr/cgi-bin/nph-acedb3.1/acedb/metabolisme (this web site provides information on *C. elegans* metabolic enzymes); http://www.gwu.edu/~mpb (this web site provides information on metabolic pathways); http://www.bic.nus.edu.sg/ pathwaydb.html (this web site provides links to biological pathways, such as metabolic pathways, developmental pathways, signal-transduction pathways, and genetic regulatory circuits); and http://www.scri.sari.ac.uk/bpp/charttxt.htm (this web site provides graphics of the metabolic pathways of diseased potato).

SUMMARY OF THE INVENTION

The present invention provides a substantially purified nucleic acid molecule where the nucleic acid molecule comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either.

The present invention provides a substantially purified first nucleic acid molecule, wherein the first nucleic acid molecule specifically hybridizes to a second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof.

The present invention provides a marker nucleic acid molecule capable of detecting the level, pattern, occurrence or absence of a biochemical process, wherein the biochemical process is selected from the group consisting of photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, lipid metabolism, biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate metabolism, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine metabolism, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone

30

5

10



metabolism, phenylpropanoid metabolism, isoprenoid metabolism, alpha-oxidation lipid metabolism, fatty acid metabolism, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

The present invention also provides a substantially purified protein or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO:463,173.

The present invention also provides a substantially purified protein or fragment thereof encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO:463,173.

The present invention also provides a purified antibody or fragment thereof which is capable of specifically binding to a protein or fragment thereof, wherein the protein or fragment thereof is encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173; and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; which is linked to (B) a nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein the transcribed strand is complementary to a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or fragment thereof; which is linked to (C) a 3' non-translated sequence that functions in plant cells to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

30

5

10



The present invention provides a microarray comprising a collection of nucleic acid molecules wherein the collection of nucleic acid molecules are capable of detecting or predicting a component or attribute of a biochemical process or activity, where the biochemical process or activity are selected from the group consisting of photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, lipid metabolism, biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate metabolism, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine metabolism, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, α-oxidation lipid metabolism, fatty acid metabolism, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

The present invention also provides a method for determining a level or pattern of a plant protein in a plant cell or plant tissue comprising: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragment of either, with a complementary nucleic acid molecule obtained from the plant cell or plant tissue, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue permits the detection of the protein; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue; and (C) detecting the level or pattern of the complementary nucleic acid, wherein the detection of the complementary nucleic acid is predictive of the level or pattern of the protein.



The present invention also provides a method for determining a level or pattern of a protein in a plant cell or plant tissue under evaluation which comprises assaying the concentration of a molecule, whose concentration is dependent upon the expression of a gene, the gene specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO: 463,173, in comparison to the concentration of that molecule present in a reference plant cell or a reference plant tissue with a known level or pattern of the protein, wherein the assayed concentration of the molecule is compared to the assayed concentration of the molecule in the reference plant cell or reference plant tissue with the known level or pattern of the protein.

The present invention provides a method of determining a mutation in a plant whose presence is predictive of a mutation affecting a level or pattern of a protein comprising the steps: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid selected from the group of marker nucleic acid molecules which specifically hybridize to a nucleic acid molecule having a nucleic acid sequence selected from the group of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof and a complementary nucleic acid molecule obtained from the plant, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant permits the detection of a polymorphism whose presence is predictive of a mutation affecting the level or pattern of the protein in the plant; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant; and (C) detecting the presence of the polymorphism, wherein the detection of the polymorphism is predictive of the mutation.

The present invention also provides a method of producing a plant containing an overexpressed protein comprising: (A) transforming the plant with a functional nucleic acid molecule,
wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter
region is linked to a structural region, wherein the structural region has a nucleic acid sequence
selected from group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173; wherein the
structural region is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA
molecule; and wherein the functional nucleic acid molecule results in overexpression of the
protein; and (B) growing the transformed plant.

30

5

10



The present invention also provides a method of producing a plant containing reduced levels of a protein comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter region is linked to a structural region, wherein the structural region comprises a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO: 463,173 or fragment thereof and the transcribed strand is complementary to an endogenous mRNA molecule; and wherein the transcribed nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and (B) growing the transformed plant.

The present invention also provides a method of determining an association between a polymorphism and a plant trait comprising: (A) hybridizing a nucleic acid molecule specific for the polymorphism to genetic material of a plant, wherein the nucleic acid molecule has a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragment of either; and (B) calculating the degree of association between the polymorphism and the plant trait.

The present invention also provides a method of isolating a nucleic acid comprising: (A) incubating under conditions permitting nucleic acid hybridization, a first nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragment of either with a complementary second nucleic acid molecule obtained from a plant cell or plant tissue; (B) permitting hybridization between the first nucleic acid molecule and the second nucleic acid molecule obtained from the plant cell or plant tissue; and (C) isolating the second nucleic acid molecule.

The present invention also provides a method of analyzing the differences in the RNA profiles from more than one physiological source, the method comprising: a) obtaining a sample of ribonucleic acids from each of the physiological sources; b) generating a population of labeled nucleic acids for each of the physiological sources from said sample of ribonucleic acids; c) hybridizing the labeled nucleic acids for each of the physiological sources to an array of nucleic acid molecules stably associated with the surface of a substrate to produce a hybridization pattern for each of the physiological sources; said stably associated nucleic acid molecules selected from

30

5

10



the group consisting of SEQ ID NO: 1 through SEQ ID NO:463,173 or fragments thereof and d) comparing the hybridization patterns for each of the different physiological sources.

DETAILED DESCRIPTION OF THE INVENTION

One skilled in the art can refer to general reference texts for detailed descriptions of known techniques discussed herein or equivalent techniques. These texts include Current Protocols in Molecular Biology Ausubel, et al., eds., John Wiley & Sons, N. Y. (1989), and supplements through September (1998), Molecular Cloning, A Laboratory Manual, Sambrook et al, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), Genome Analysis: A Laboratory Manual 1: Analyzing DNA, Birren et al., Cold Spring Harbor Press, Cold Spring Harbor, New York (1997); Genome Analysis: A Laboratory Manual 2: Detecting Genes, Birren et al., Cold Spring Harbor Press, Cold Spring Harbor, New York (1998); Genome Analysis: A Laboratory Manual 3: Cloning Systems, Birren et al., Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); Genome Analysis: A Laboratory Manual 4: Mapping Genomes, Birren et al., Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); Plant Molecular Biology: A Laboratory Manual, Clark, Springer-Verlag, Berlin, (1997), Methods in Plant Molecular Biology, Maliga et al., Cold Spring Harbor Press, Cold Spring Harbor, New York (1995). These texts can, of course, also be referred to in making or using an aspect of the invention. It is understood that any of the agents of the invention can be substantially purified and/or be biologically active and/or recombinant.

Agents

(a) Nucleic Acid Molecules

Agents of the present invention include plant nucleic acid molecules and more preferably include maize, soybean, cotton, sorghum, teosinte, wheat, and rice nucleic acid molecules.

A subset of the nucleic acid molecules of the present invention includes nucleic acid molecules that are marker molecules. Another subset of the nucleic acid molecules of the present invention includes nucleic acid molecules that encode a protein or fragment thereof. Another subset of the nucleic acid molecules of the present invention is cDNA molecules.

Fragment nucleic acid molecules may encode significant portion(s) of, or indeed most of, these nucleic acid molecules. Alternatively, the fragments may comprise smaller oligonucleotides (having from about 15 to about 250 nucleotide residues and more preferably, about 15 to

30

5

10



about 30 nucleotide residues, or more preferably about 30 to about 50 nucleotide residues, or again more preferably about 50 to about 100 nucleotide residues).

The term "substantially purified," as used herein, refers to a molecule separated from substantially all other molecules normally associated with it in its native state. More preferably a substantially purified molecule is the predominant species present in a preparation. A substantially purified molecule may be greater than 60% free, preferably 75% free, more preferably 90% free, and most preferably 95% free from the other molecules (exclusive of solvent) present in the natural mixture. The term "substantially purified" is not intended to encompass molecules present in their native state.

The agents of the present invention will preferably be "biologically active" with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid molecule, or the ability of a protein to be bound by an antibody (or to compete with another molecule for such binding). Alternatively, such an attribute may be catalytic and thus involve the capacity of the agent to mediate a chemical reaction or response.

The agents of the present invention may also be recombinant. As used herein, the term recombinant, refers to a) molecules that are constructed outside of living cells by joining natural or synthetic DNA segments to DNA molecules that can replicate in a living cell or b) molecules that result from the replication or expression of those molecules described above or c) amino acid molecules from different sources which are joined together.

It is understood that the agents of the present invention may be labeled with reagents that facilitate detection of the agent (e.g., fluorescent labels, Prober et al., Science 238:336-340 (1987); Albarella et al., EP 144914; chemical labels, Sheldon et al., U.S. Patent 4,582,789; Albarella et al., U.S. Patent 4,563,417; modified bases, Miyoshi et al., EP 119448).

It is further understood, that the present invention provides recombinant bacterial, mammalian, microbial, insect, fungal and plant cells and viral constructs comprising the agents of the present invention

Nucleic acid molecules or fragments thereof of the present invention are capable of specifically hybridizing to other nucleic acid molecules under certain circumstances. As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure. A

5

10



nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if they exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are described by Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989) and by Haymes et al., Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, DC (1985). Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for a nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 50°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 X SSC at 50°C to a high stringency of about 0.2 X SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof under moderately stringent conditions, for example at about 2.0 X SSC and about 65°C.

30

5

10

In a particularly preferred embodiment, a nucleic acid of the present invention will include those nucleic acid molecules that specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof under high stringency conditions such as 0.2 X SSC and about 65°C.

In one aspect of the present invention, the nucleic acid molecules of the present invention comprise one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either. In another aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 90% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 95% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either. In a more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 98% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 complements thereof or fragments of either. In an even more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 99% sequence identity with one or more of the sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof.

The term "sequence identity" refers to the extent to which two sequences, nucleotide or amino acid, are invariant throughout the portion at which they are aligned. While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "sequence identity" is well known to skilled artisans. Methods commonly employed to determine identity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., SIAM J Applied Math (1988) 48:1073. Methods to determine identity are codified in computer programs. Preferred computer program methods to determine identity between two sequences include, but are not limited to, the BLAST suite of programs publicly available from NCBI and other sources (BLAST Manual, Altschul et al., Natl. Cent. Biotechnol. Inf.,

30

5

10

Natl. Library Med. (NCBI NLM) NIH, Bethesda, Md. 20894; Altschul et al., J. Mol. Biol. 215:403-410 (1990), Pearson et al., Proc. Natl. Acad. Sci. U.S.A. 85:2444-2448 (1988), the FAST programs (Pearson et al., Proc. Natl. Acad. Sci. U.S.A. 85:2444-2448 (1988).the GAP and BESTFIT programs found in the GCG program package, (Madison, WI) and Cross_Match (Phi Green, University of Washington). Another preferred method to determine identity, is by the method of DNASTAR protein alignment protocol using the Jotun-Hein algorithm (Hein et al., Methods Enzymol. 183:626-645 (1990)).

Unless otherwise noted, "percent sequence identity or percent identity" for this invention refers to the value obtained when using the BLAST 2.0 suite of programs with default parameters (Altschul et al., Nucleic Acids Res. 25:3389-3402, 1997; Altschul et al., J. Mol. Bio. 215: 403-410, 1990) Version 2.0 of BLAST allows the introduction of gaps (deletions and insertions) into alignments.

(i) Nucleic Acid Molecules Encoding Proteins or Fragments Thereof

Nucleic acid molecules of the present invention can comprise sequences that encode a protein or fragment thereof. Such proteins or fragments thereof include homologues of known proteins in other organisms.

In a preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of another plant protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a fungal protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein of the present invention is a homologue of a mammalian protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a bacterial protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a maize protein. In another preferred embodiment of the present invention, a soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a soybean protein. In another preferred embodiment of the present invention is a homologue of a soybean protein. In

5

10



cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a cotton protein. In another preferred embodiment of the present invention, a maize, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a wheat protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of an *Arabidopsis* protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof of the present invention is a homologue of a sorghum protein. In another preferred embodiment of the present invention, a maize, soybean, *Arabidopsis*, wheat, cotton, or rice protein or fragment thereof of the present invention is a homologue of a teosinte protein.

In a preferred embodiment of the present invention, the nucleic molecule of the present invention encodes a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein or fragment thereof where a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein exhibits a BLAST E value score of greater than 1E-12, preferably a BLAST E value score of between about 1E-30 and about 1E-12, even more preferably a BLAST probability E value score of greater than 1E-30 with its homologue.

Nucleic acid molecules of the present invention also include non-maize, non-soybean, non-rice, non-wheat, non-Arabidopsis, non-sorghum, non-cotton and non-teosinte homologues. Preferred plant sources of homologues are selected from the group consisting of alfalfa, barley, Brassica, broccoli, cabbage, citrus, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rice, rye, strawberry, sugarcane, sugarbeet, tomato, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm and Phaseolus.

In a preferred embodiment, nucleic acid molecules having SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements and fragments of either can be utilized to obtain such homologues.

In another further aspect of the present invention, nucleic acid molecules of the present invention can comprise sequences which differ from those encoding a protein or fragment thereof in SEQ ID NO: 1 through SEQ ID NO: 463,173 due to fact that the different nucleic acid sequence

30

5

10

encodes a protein having one or more conservative amino acid changes. It is understood that codons capable of coding for such conservative amino acid substitutions are known in the art.

It is well known in the art that one or more amino acids in a native sequence can be substituted with another amino acid(s), the charge and polarity of which are similar to that of the native amino acid, *i.e.*, a conservative amino acid substitution, resulting in a silent change. Conserved substitutions for an amino acid within the native polypeptide sequence can be selected from other members of the class to which the naturally occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids, (2) basic amino acids, (3) neutral polar amino acids, and (4) neutral nonpolar amino acids. Representative amino acids within these various groups include, but are not limited to, (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cysteine, cystine, tyrosine, asparagine, and glutamine; and (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine.

Conservative amino acid changes within the native polypeptides sequence can be made by substituting one amino acid within one of these groups with another amino acid within the same group. Biologically functional equivalents of the proteins or fragments thereof of the present invention can have ten or fewer conservative amino acid changes, more preferably seven or fewer conservative amino acid changes, and most preferably five or fewer conservative amino acid changes. The encoding nucleotide sequence will thus have corresponding base substitutions, permitting it to encode biologically functional equivalent forms of the proteins or fragments of the present invention.

It is understood that certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Because it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence and, of course, its underlying DNA coding sequence and, nevertheless, obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the pep-

30

5

10



tide sequences of the proteins or fragments of the present invention, or corresponding DNA sequences that encode said peptides, without appreciable loss of their biological utility or activity. It is understood that codons capable of coding for such amino acid changes are known in the art.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte and Doolittle, *J. Mol. Biol. 157*, 105-132 (1982)). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like.

Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, *J. Mol. Biol. 157*, 105-132 (1982)); these are isoleucine (+4.5), valine (+4.2), leucine (+3.8), phenylalanine (+2.8), cysteine/cystine (+2.5), methionine (+1.9), alanine (+1.8), glycine (-0.4), threonine (-0.7), serine (-0.8), tryptophan (-0.9), tyrosine (-1.3), proline (-1.6), histidine (-3.2), glutamate (-3.5), glutamine (-3.5), asparate (-3.5), asparagine (-3.5), lysine (-3.9), and arginine (-4.5).

In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those that are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101 states that the greatest local average hydrophilicity of a protein, as govern by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0), lysine (+3.0), aspartate (+3.0 \pm 1), glutamate (+3.0 \pm 1), serine (+0.3), asparagine (+0.2), glutamine (+0.2), glycine (0), threonine (-0.4), proline (-0.5 \pm 1), alanine (-0.5), histidine (-0.5), cysteine (-1.0), methionine (-1.3), valine (-1.5), leucine (-1.8), isoleucine (-1.8), tyrosine (-2.3), phenylalanine (-2.5), and tryptophan (-3.4).

In making such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those that are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

30

5

10



In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those encoding a protein or fragment thereof set forth in SEQ ID NO: 1 through SEQ ID NO: 463,173 or fragment thereof due to the fact that one or more codons encoding an amino acid has been substituted for a codon that encodes a nonessential substitution of the amino acid originally encoded.

Agents of the invention include nucleic acid molecules that encode at least about a contiguous 10 amino acid region of a protein of the present invention, more preferably at least about a contiguous 25, 40, 50, 100, or 125 amino acid region of a protein of the present invention. In a preferred embodiment the protein is selected from the group consisting of a plant, more preferably a maize, soybean, *Arabidopsis*, wheat, cotton, teosinte, or rice protein.

(ii) Nucleic Acid Molecule Markers and Probes

One aspect of the present invention concerns nucleic acid molecules of the present invention that can act as markers, for example, those nucleic acid molecules SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either that can act as markers or one or more of the marker molecules encoded by other nucleic acid agents of the present invention.

As used herein, a "marker" is an indicator for the presence of at least one polymorphism. A marker is preferably a nucleic acid molecule.

A "nucleic acid marker" as used herein means a nucleic acid molecule that is capable of being a marker for detecting a polymorphism.

In a preferred embodiment, the level, pattern, occurrence and/or absence of a nucleic acid molecule and/or collection of nucleic acid molecules of the present invention is a marker, for example, for a developmental, commercial or non-commercially valuable trait such as yield or an environmental condition or treatment. It is noted that many agronomic traits can affect yield. These include, without limitation, pod position on the plant, number of internodes, incidence of pod shatter, grain size, efficiency of nodulation and nitrogen fixation, efficiency of nutrient assimilation, resistance to biotic and abiotic stress, carbon assimilation, plant architecture, resistance to lodging, percent seed germination, seedling vigor, and juvenile traits.

As used herein, a "collection of nucleic acid molecules" is a population of nucleic acid molecules where at least two of the nucleic acid molecules differ, at least in part, in their nucleic acid sequence. It is understood, that as used herein, an individual species within a collection of

5

10

nucleic acid molecules may be physically separate or alternatively not physically separate from one or more other species within the collection of nucleic acid molecules. An example of a situation where individual species may be physically separate but considered a collection of nucleic acid molecules is where more than two species are present on a single support such as a nylon membrane or a glass but occupy a different position on such support. Examples of situations where individual species are physically separate on a support include microarrays.

As used herein, where a collection of nucleic acids is a marker for a particular attribute, the level, pattern, occurrence and/or absence of the nucleic acid molecules associated with the attribute are not required to be the same between species of the collection. For example, the increase in the level of a species when in combination with the decrease in a second species could be diagnostic for a particular attribute.

In an even more preferred embodiment of the present invention, the level, pattern, occurrence and/or absence of a nucleic acid molecule and/or collection of nucleic acid molecules of the present invention is a marker for a biochemical process or activity where the process or activity is preferably selected from photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, and lipid metabolism, and more preferably selected from the group consisting of biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis and gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate synthesis/degradation, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine metabolism, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid synthesis metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen and sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, alpha-oxidation lipid metabolism, and fatty acid metabolism, and even more preferably selected from the group consisting of: glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved

5

10



in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

Genetic markers of the invention include "dominant" or "codominant" markers. "Codominant markers" reveal the presence of two or more alleles (two per diploid individual) at a locus. "Dominant markers" reveal the presence of only a single allele per locus. The presence of the dominant marker phenotype (e.g., a band of DNA) is an indication that one allele is in either the homozygous or heterozygous condition. The absence of the dominant marker phenotype (e.g., absence of a DNA band) is merely evidence that "some other" undefined allele is present. In the case of populations where individuals are predominantly homozygous and loci are predominately dimorphic, dominant and codominant markers can be equally valuable. As populations become more heterozygous and multi-allelic, codominant markers often become more informative of the genotype than dominant markers. Marker molecules can be, for example, capable of detecting polymorphisms such as single nucleotide polymorphisms (SNPs).

SNPs can be characterized using any of a variety of methods (Botstein et al., Am. J. Hum. Genet. 32:314-331 (1980); Konieczny and Ausubel, Plant J. 4:403-410 (1993); Myers et al., Nature 313:495-498 (1985); Newton et al., Nucl. Acids Res. 17:2503-2516 (1989); Wu et al., Proc. Natl. Acad. Sci. (U.S.A.) 86:2757-2760 (1989); Barany, Proc. Natl. Acad. Sci. (U.S.A.) 88:189-193 (1991); Labrune et al., Am. J. Hum. Genet. 48: 1115-1120 (1991); Kuppuswami et al., Proc. Natl. Acad. Sci. USA 88:1143-1147 (1991); Sarkar et al., Genomics 13:441-443 (1992); Nikiforov et al., Nucl. Acids Res. 22:4167-4175 (1994); Livak et al., PCR Methods Appl. 4:357-362 (1995); Livak et al., Nature Genet. 9:341-342 (1995); Chen and Kwok, Nucl. Acids Res. 25:347-353 (1997); Tyagi et al., Nature Biotech. 16: 49-53 (1998); Haff and Smirnov, Genome Res. 7: 378-388 (1997); Neff et al., Plant J. 14:387-392 (1998)).

Additional markers, such as AFLP markers, RFLP markers and RAPD markers, can be utilized (Walton, *Seed World* 22-29 (July, 1993); Burow and Blake, *Molecular Dissection of Complex Traits*, 13-29, Paterson (ed.), CRC Press, New York (1988)). Another marker type, RAPDs, is developed from DNA amplification with random primers and result from single base changes and insertions/deletions in plant genomes. They are dominant markers with a medium level of polymorphisms and are highly abundant. AFLP markers require using the PCR on a

30

5

10

subset of restriction fragments from extended adapter primers. These markers are both dominant and codominant are highly abundant in genomes and exhibit a medium level of polymorphism.

A PCR probe is a nucleic acid molecule capable of initiating a polymerase activity while in a double-stranded structure to with another nucleic acid. Various methods for determining the structure of PCR probes and PCR techniques exist in the art. Computer generated searches using programs such as Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www-genome.wi.mit.edu/cgi-bin/www-STS_Pipeline), or GeneUp (Pesole *et al.*, *BioTechniques 25*:112-123 (1998)), for example, can be used to identify potential PCR primers.

It is understood that a fragment of one or more of the nucleic acid molecules of the present invention may be a probe and preferably a PCR probe.

(b) Protein and Peptide Molecules

A class of agents comprises one or more of the protein or peptide molecules encoded by SEQ ID NO: 1 through SEQ ID NO: 463,173 or one or more of the protein or fragment thereof or peptide molecules encoded by other nucleic acid agents of the present invention. As used herein, the term "protein molecule" or "peptide molecule" includes any molecule that comprises five or more amino acids. It is well know in the art that proteins may undergo modification, including post-translational modifications, such as, but not limited to, disulfide bond formation, glycosylation, phosphorylation, or oligomerization. Thus, as used herein, the term "protein molecule" or "peptide molecule" includes any protein molecule that is modified by any biological or non-biological process. The terms "amino acid" and "amino acids" refer to all naturally occurring L-amino acids. This definition is meant to include norleucine, ornithine, homocysteine, and homoserine.

One or more of the protein or fragment of peptide molecules may be produced via chemical synthesis, or more preferably, by expression in a suitable bacterial or eukaryotic host. Suitable methods for expression are described by Sambrook, et al., (In: Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989)), or similar texts.

A "protein fragment" is a peptide or polypeptide molecule whose amino acid sequence comprises a subset of the amino acid sequence of that protein. A protein or fragment thereof that comprises one or more additional peptide regions not derived from that protein is a "fusion"

30

5

10

protein. Such molecules may be derivatized to contain carbohydrate or other moieties (such as keyhole limpet hemocyanin, etc.). Fusion protein or peptide molecules of the present invention are preferably produced via recombinant means.

Another class of agents comprise protein or peptide molecules encoded by SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or, fragments or fusions thereof in which non-essential, or not relevant, amino acid residues have been added, replaced, or deleted. An example of such a homologue is the homologue protein of all non- soybean, non-maize, non-rice, non-cotton, non-sorghum, non-teosinte, non-Arabidopsis and non-wheat plant species, including but not limited to alfalfa, barley, Brassica, broccoli, cabbage, citrus, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rye, strawberry, sugarcane, sugarbeet, tomato, poplar, pine, fir, eucalyptus, apple, lettuce, peas, lentils, grape, banana, tea, turf grasses, etc. Particularly preferred non-soybean, non-maize, non-rice, non-cotton, non-sorghum, non-teosinte, non-Arabidopsis and non-wheat plants to utilize for the isolation of homologues would include alfalfa, barley, oat, oilseed rape, canola, ornamentals, sugarcane, sugarbeet, tomato, potato, and turf grasses. Such a homologue can be obtained by any of a variety of methods. Most preferably, as indicated above, one or more of the disclosed sequences (SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof) will be used to define a pair of primers that may be used to isolate the homologue-encoding nucleic acid molecules from any desired species. Such molecules can be expressed to yield homologues by recombinant means.

(c) Antibodies

One aspect of the present invention concerns antibodies, single-chain antigen binding molecules, or other proteins that specifically bind to one or more of the protein or peptide molecules of the present invention and their homologues, fusions or fragments. Such antibodies may be used to quantitatively or qualitatively detect the protein or peptide molecules of the present invention. As used herein, an antibody or peptide is said to "specifically bind" to a protein or peptide molecule of the present invention if such binding is not competitively inhibited by the presence of non-related molecules.

Nucleic acid molecules that encode all or part of the protein of the present invention can be expressed, via recombinant means, to yield protein or peptides that can in turn be used to elicit antibodies that are capable of binding the expressed protein or peptide. Such antibodies

30

5

10

may be used in immunoassays for that protein. Such protein-encoding molecules, or their fragments may be a "fusion" molecule (*i.e.*, a part of a larger nucleic acid molecule) such that, upon expression, a fusion protein is produced. It is understood that any of the nucleic acid molecules of the present invention may be expressed, via recombinant means, to yield proteins or peptides encoded by these nucleic acid molecules.

The antibodies that specifically bind proteins and protein fragments of the present invention may be polyclonal or monoclonal and may comprise intact immunoglobulins, or antigen binding portions of immunoglobulins fragments (such as (F(ab'), F(ab')₂), or single-chain immunoglobulins producible, for example, via recombinant means. It is understood that practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (see, for example, Harlow and Lane, In: Antibodies: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, New York (1988)).

Murine monoclonal antibodies are particularly preferred. BALB/c mice are preferred for this purpose, however, equivalent strains may also be used. The animals are preferably immunized with approximately 25 µg of purified protein (or fragment thereof) that has been emulsified in a suitable adjuvant (such as TiterMax adjuvant (Vaxcel, Norcross, GA)). Immunization is preferably conducted at two intramuscular sites, one intraperitoneal site and one subcutaneous site at the base of the tail. An additional i.v. injection of approximately 25 µg of antigen is preferably given in normal saline three weeks later. After approximately 11 days following the second injection, the mice may be bled and the blood screened for the presence of anti-protein or peptide antibodies. Preferably, a direct binding Enzyme-Linked Immunoassay (ELISA) is employed for this purpose.

More preferably, the mouse having the highest antibody titer is given a third i.v. injection of approximately 25 µg of the same protein or fragment. The splenic leukocytes from this animal may be recovered 3 days later and then permitted to fuse, most preferably, using polyethylene glycol, with cells of a suitable myeloma cell line (such as, for example, the P3X63Ag8.653 myeloma cell line). Hybridoma cells are selected by culturing the cells under "HAT" (hypoxanthine-aminopterin-thymine) selection for about one week. The resulting clones may then be screened for their capacity to produce monoclonal antibodies ("mAbs"), preferably by direct ELISA.

30

5

10

In one embodiment, anti-protein or peptide monoclonal antibodies are isolated using a fusion of a protein or peptide of the present invention, or conjugate of a protein or peptide of the present invention, as immunogens. Thus, for example, a group of mice can be immunized using a fusion protein emulsified in Freund's complete adjuvant (e.g. approximately 50 µg of antigen per immunization). At three-week intervals, an identical amount of antigen is emulsified in Freund's incomplete adjuvant and used to immunize the animals. Ten days following the third immunization, serum samples are taken and evaluated for the presence of antibody. If antibody titers are too low, a fourth booster can be employed. Polysera capable of binding the protein or peptide can also be obtained using this method.

In a preferred procedure for obtaining monoclonal antibodies, the spleens of the above-described immunized mice are removed, disrupted and immune splenocytes are isolated over a ficoll gradient. The isolated splenocytes are fused, using polyethylene glycol with BALB/c-derived HGPRT (hypoxanthine guanine phosphoribosyl transferase) deficient P3x63xAg8.653 plasmacytoma cells. The fused cells are plated into 96 well microtiter plates and screened for hybridoma fusion cells by their capacity to grow in culture medium supplemented with hypothanthine, aminopterin and thymidine for approximately 2-3 weeks.

Hybridoma cells that arise from such incubation are preferably screened for their capacity to produce an immunoglobulin that binds to a protein of interest. An indirect ELISA may be used for this purpose. In brief, the supernatants of hybridomas are incubated in microtiter wells that contain immobilized protein. After washing, the titer of bound immunoglobulin can be determined using, for example, a goat anti-mouse antibody conjugated to horseradish peroxidase. After additional washing, the amount of immobilized enzyme is determined (for example through the use of a chromogenic substrate). Such screening is performed as quickly as possible after the identification of the hybridoma in order to ensure that a desired clone is not overgrown by non-secreting neighbor cells. Desirably, the fusion plates are screened several times since the rates of hybridoma growth vary. In a preferred sub-embodiment, a different antigenic form may be used to screen the hybridoma. Thus, for example, the splenocytes may be immunized with one immunogen, but the resulting hybridomas can be screened using a different immunogen. It is understood that any of the protein or peptide molecules of the present invention may be used to raise antibodies.

5

10

As discussed below, such antibody molecules or their fragments may be used for diagnostic purposes. Where the antibodies are intended for diagnostic purposes, it may be desirable to derivatize them, for example with a ligand group (such as biotin) or a detectable marker group (such as a fluorescent group, a radioisotope or an enzyme).

The ability to produce antibodies that bind the protein or peptide molecules of the present invention permits the identification of mimetic compounds of those molecules. A "mimetic compound" is a compound that is not that compound, or a fragment of that compound, but which nonetheless exhibits an ability to specifically bind to antibodies directed against that compound.

It is understood that any of the agents of the present invention can be substantially purified and/or be biologically active and/or recombinant.

(d) Plant Constructs and Plant Transformants

One or more of the nucleic acid molecules of the invention may be used in plant transformation or transfection. Exogenous genetic material may be transferred into a plant cell and the plant cell regenerated into a whole, fertile or sterile plant. Exogenous genetic material is any genetic material, whether naturally occurring or otherwise, from any source that is capable of being inserted into any organism. In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention, preferably a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either.

Such genetic material may be transferred into either monocotyledons and dicotyledons including, but not limited to maize, soybean, *Arabidopsis*, phaseolus, peanut, alfalfa, wheat, rice, oat, sorghum, rye, tritordeum, millet, fescue, perennial ryegrass, sugarcane, cranberry, papaya, banana, banana, muskmelon, apple, cucumber, dendrobium, gladiolus, chrysanthemum, liliacea, cotton, eucalyptus, sunflower, canola, turfgrass, sugarbeet, coffee and dioscorea (Christou, In: *Particle Bombardment for Genetic Engineering of Plants*, Biotechnology Intelligence Unit. Academic Press, San Diego, California (1996)).

Transfer of a nucleic acid that encodes for a protein can result in overexpression of that protein in a transformed cell or transgenic plant. One or more of the proteins or fragments thereof encoded by nucleic acid molecules of the invention may be overexpressed in a

30

5

10

transformed cell or transformed plant. Such overexpression may be the result of transient or stable transfer of the exogenous genetic material.

Exogenous genetic material may be transferred into a host cell by the use of a DNA vector or construct designed for such a purpose. Design of such a vector is generally within the skill of the art (*See*, *Plant Molecular Biology: A Laboratory Manual*, Clark (ed.), Springier, New York (1997)).

A construct or vector may include a plant promoter to express the protein or protein fragment of choice. A number of promoters, which are active in plant cells, have been described in the literature. These include the nopaline synthase (NOS) promoter (Ebert et al., Proc. Natl. Acad. Sci. (U.S.A.) 84:5745-5749 (1987)), the octopine synthase (OCS) promoter (which are carried on tumor-inducing plasmids of Agrobacterium tumefaciens), the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton et al., Plant Mol. Biol. 9:315-324 (1987)) and the CaMV 35S promoter (Odell et al., Nature 313:810-812 (1985)), the figwort mosaic virus 35S-promoter, the light-inducible promoter from the small subunit of ribulose-1,5-bis-phosphate carboxylase (ssRUBISCO), the Adh promoter (Walker et al., Proc. Natl. Acad. Sci. (U.S.A.) 84:6624-6628 (1987)), the sucrose synthase promoter (Yang et al., Proc. Natl. Acad. Sci. (U.S.A.) 87:4144-4148 (1990)), the R gene complex promoter (Chandler et al., The Plant Cell 1:1175-1183 (1989)) and the chlorophyll a/b binding protein gene promoter, etc. These promoters have been used to create DNA constructs that have been expressed in plants; see, e.g., PCT publication WO 84/02913. The CaMV 35S promoters are preferred for use in plants. Promoters known or found to cause transcription of DNA in plant cells can be used in the invention.

For the purpose of expression in source tissues of the plant, such as the leaf, seed, root or stem, it is preferred that the promoters utilized have relatively high expression in these specific tissues. Tissue-specific expression of a protein of the present invention is a particularly preferred embodiment. For this purpose, one may choose from a number of promoters for genes with tissue- or cell-specific or -enhanced expression. Examples of such promoters reported in the literature include the chloroplast glutamine synthetase GS2 promoter from pea (Edwards *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:3459-3463 (1990)), the chloroplast fructose-1,6-biphosphatase (FBPase) promoter from wheat (Lloyd *et al.*, *Mol. Gen. Genet.* 225:209-216 (1991)), the nuclear

30

5

10



photosynthetic ST-LS1 promoter from potato (Stockhaus et al., EMBO J. 8:2445-2451 (1989)), the serine/threonine kinase (PAL) promoter and the glucoamylase (CHS) promoter from Arabidopsis thaliana. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase (RbcS) promoter from eastern larch (*Larix laricina*), the promoter for the cab gene, cab6, from pine (Yamamoto et al., Plant Cell Physiol. 35:773-778 (1994)), the promoter for the Cab-1 gene from wheat (Fejes et al., Plant Mol. Biol. 15:921-932 (1990)), the promoter for the CAB-1 gene from spinach (Lubberstedt et al., Plant Physiol. 104:997-1006 (1994)), the promoter for the cab1R gene from rice (Luan et al., Plant Cell. 4:971-981 (1992)), the pyruvate, orthophosphate dikinase (PPDK) promoter from maize (Matsuoka et al., Proc. Natl. Acad. Sci. (U.S.A.) 90: 9586-9590 (1993)), the promoter for the tobacco Lhcb1*2 gene (Cerdan et al., Plant Mol. Biol. 33:245-255 (1997)), the Arabidopsis thaliana SUC2 sucrose-H+ symporter promoter (Truernit et al., Planta. 196:564-570 (1995)) and the promoter for the thylakoid membrane proteins from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Other promoters for the chlorophyll a/b-binding proteins may also be utilized in the invention, such as the promoters for LhcB gene and PsbP gene from white mustard (Sinapis alba; Kretsch et al., Plant Mol. Biol. 28:219-229 (1995)).

For the purpose of expression in sink tissues of the plant, such as the tuber of the potato plant, the fruit of tomato, or the seed of maize, wheat, rice and barley, it is preferred that the promoters utilized in the invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or -enhanced expression are known, including the class I patatin promoter (Bevan *et al.*, *EMBO J. 8*:1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol. 14*:995-1006 (1990)), the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter (Salanoubat and Belliard, *Gene 60*:47-56 (1987), Salanoubat and Belliard, *Gene 84*:181-185 (1989)), the promoter for the major tuber proteins including the 22 kd protein complexes and proteinase inhibitors (Hannapel, *Plant Physiol. 101*:703-704 (1993)), the promoter for the granule bound starch synthase gene (GBSS) (Visser *et al.*, *Plant Mol. Biol. 17*:691-699 (1991)) and other class I and II patatins promoters (Koster-Topfer *et al.*, *Mol Gen Genet. 219*:390-396 (1989); Mignery *et al.*, *Gene. 62*:27-44 (1988)).

Other promoters can also be used to express a protein or fragment thereof in specific tissues, such as seeds or fruits. The promoter for β -conglycinin (Chen *et al.*, *Dev. Genet. 10*: 112-

30

5

10



122 (1989)) or other seed-specific promoters such as the napin and phaseolin promoters, can be used. The zeins are a group of storage proteins found in maize endosperm. Genomic clones for zein genes have been isolated (Pedersen et al., Cell 29:1015-1026 (1982)) and the promoters from these clones, including the 15 kD, 16 kD, 19 kD, 22 kD, 27 kD and genes, could also be used. Other promoters known to function, for example, in maize include the promoters for the following genes: waxy, Brittle, Shrunken 2, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins and sucrose synthases. A particularly preferred promoter for maize endosperm expression is the promoter for the glutelin gene from rice, more particularly the Osgt-1 promoter (Zheng et al., Mol. Cell Biol. 13:5829-5842 (1993)). Examples of promoters suitable for expression in wheat include those promoters for the ADPglucose pyrosynthase (ADPGPP) subunits, the granule bound and other starch synthase, the branching and debranching enzymes, the embryogenesis-abundant proteins, the gliadins and the glutenins. Examples of such promoters in rice include those promoters for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases and the glutelins. A particularly preferred promoter is the promoter for rice glutelin, Osgt-1. Examples of such promoters for barley include those for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases, the hordeins, the embryo globulins and the aleurone specific proteins.

Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene (Samac *et al.*, *Plant Mol. Biol. 25*:587-596 (1994)). Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV35S promoter that have been identified (Lam *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 86*:7890-7894 (1989)). Other root cell specific promoters include those reported by Conkling *et al.* (Conkling *et al.*, *Plant Physiol. 93*:1203-1211 (1990)).

Additional promoters that may be utilized are described, for example, in U.S. Patent Nos. 5,378,619; 5,391,725; 5,428,147; 5,447,858; 5,608,144; 5,608,144; 5,614,399; 5,633,441; 5,633,435; and 4,633,436. In addition, a tissue specific enhancer may be used (Fromm *et al.*, *The Plant Cell 1*:977-984 (1989)).

Constructs or vectors may also include, with the coding region of interest, a nucleic acid sequence that acts, in whole or in part, to terminate transcription of that region. A number of

30

5

10



such sequences have been isolated, including the Tr7 3' sequence and the NOS 3' sequence (Ingelbrecht et al., The Plant Cell 1:671-680 (1989); Bevan et al., Nucleic Acids Res. 11:369-385 (1983)).

A vector or construct may also include regulatory elements. Examples of such include the Adh intron 1 (Callis *et al.*, *Genes and Develop. 1*:1183-1200 (1987)), the sucrose synthase intron (Vasil *et al.*, *Plant Physiol. 91*:1575-1579 (1989)) and the TMV omega element (Gallie *et al.*, *The Plant Cell 1*:301-311 (1989)). These and other regulatory elements may be included when appropriate.

A vector or construct may also include a selectable marker. Selectable markers may also be used to select for plants or plant cells that contain the exogenous genetic material. Examples of such include, but are not limited to: a neo gene (Potrykus *et al.*, *Mol. Gen. Genet. 199*:183-188 (1985)), which codes for kanamycin resistance and can be selected for using kanamycin, G418, etc.; a bar gene which codes for bialaphos resistance; a mutant EPSP synthase gene (Hinchee *et al.*, *Bio/Technology 6*:915-922 (1988)) which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil (Stalker *et al.*, *J. Biol. Chem. 263*:6310-6314 (1988)); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204 (Sept. 11, 1985)); and a methotrexate resistant DHFR gene (Thillet *et al.*, *J. Biol. Chem. 263*:12500-12508 (1988)).

A vector or construct may also include a transit peptide. Incorporation of a suitable chloroplast transit peptide may also be employed (European Patent Application Publication Number 0218571). Translational enhancers may also be incorporated as part of the vector DNA. DNA constructs could contain one or more 5' non-translated leader sequences that may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the mRNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence. For a review of optimizing expression of transgenes, see Koziel *et al.*, *Plant Mol. Biol. 32*:393-405 (1996).

A vector or construct may also include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include: a β-glucuronidase or uidA gene (GUS) which encodes an enzyme for which various chromogenic substrates are known (Jeff-

30

5

10



erson, *Plant Mol. Biol, Rep. 5:*387-405 (1987); Jefferson *et al.*, *EMBO J. 6:*3901-3907 (1987)); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, Stadler Symposium *11:*263-282 (1988)); a β-lactamase gene (Sutcliffe *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 75:3737-3741 (1978)), a gene which encodes an enzyme for which various chromogenic substrates are known (e.g., PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al.*, *Science 234:*856-859 (1986)); a xylE gene (Zukowsky *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 80:1101-1105 (1983)) which encodes a catechol dioxygenase that can convert chromogenic catechols; an α-amylase gene (Ikatu *et al.*, *Bio/Technol.* 8:241-242 (1990)); a tyrosinase gene (Katz *et al.*, *J. Gen. Microbiol. 129:*2703-2714 (1983)) which encodes an enzyme capable of oxidizing tyrosine to DOPA and dopaquinone which in turn condenses to melanin; an α-galactosidase, which will turn a chromogenic α-galactose substrate.

Included within the terms "selectable or screenable marker genes" are also genes that encode a secretable marker whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers that encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes that can be detected catalytically. Secretable proteins fall into a number of classes, including small, diffusible proteins which are detectable, (e.g., by ELISA), small active enzymes which are detectable in extracellular solution (e.g., α -amylase, β -lactamase, phosphinothricin transferase), or proteins which are inserted or trapped in the cell wall (such as proteins which include a leader sequence such as that found in the expression unit of extension or tobacco PR-S). Other possible selectable and/or screenable marker genes will be apparent to those of skill in the art.

There are many methods for introducing transforming nucleic acid molecules into plant cells. Suitable methods are believed to include virtually any method by which nucleic acid molecules may be introduced into a cell, such as by *Agrobacterium* infection or direct delivery of nucleic acid molecules such as, for example, by PEG-mediated transformation, by electroporation or by acceleration of DNA coated particles, etc (Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225 (1991); Vasil, *Plant Mol. Biol.* 25:925-937 (1994)). For example, electroporation has been used to transform maize protoplasts (Fromm *et al.*, *Nature* 312:791-793 (1986)).

Other vector systems suitable for introducing transforming DNA into a host plant cell include but are not limited to binary artificial chromosome (BIBAC) vectors (Hamilton *et al.*,

30

5

10



Gene 200:107-116 (1997)); and transfection with RNA viral vectors (Della-Cioppa et al., Ann. N.Y. Acad. Sci. (1996), 792 (Engineering Plants for Commercial Products and Applications), 57-61). Additional vector systems also include plant selectable YAC vectors such as those described in Mullen et al., Molecular Breeding 4:449-457 (1988)).

Technology for introduction of DNA into cells is well known to those of skill in the art. Four general methods for delivering a gene into cells have been described: (1) chemical methods (Graham and van der Eb, *Virology* 54:536-539 (1973)); (2) physical methods such as microinjection (Capecchi, *Cell* 22:479-488 (1980)), electroporation (Wong and Neumann, *Biochem. Biophys. Res. Commun.* 107:584-587 (1982); Fromm *et al.*, *Proc. Natl. Acad. Sci.* (U.S.A.) 82:5824-5828 (1985); U.S. Patent No. 5,384,253); and the gene gun (Johnston and Tang, *Methods Cell Biol.* 43:353-365 (1994)); (3) viral vectors (Clapp, *Clin. Perinatol.* 20:155-168 (1993); Lu *et al.*, *J. Exp. Med.* 178:2089-2096 (1993); Eglitis and Anderson, *Biotechniques* 6:608-614 (1988)); and (4) receptor-mediated mechanisms (Curiel *et al.*, *Hum. Gen. Ther.* 3:147-154 (1992), Wagner *et al.*, *Proc. Natl. Acad. Sci.* (USA) 89:6099-6103 (1992)).

Acceleration methods that may be used include, for example, microprojectile bombard-ment and the like. One example of a method for delivering transforming nucleic acid molecules to plant cells is microprojectile bombardment. This method has been reviewed by Yang and Christou (eds.), *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994)). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum and the like.

A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly transforming monocots, is that neither the isolation of protoplasts (Cristou *et al.*, *Plant Physiol.* 87:671-674 (1988)) nor the susceptibility of *Agrobacterium* infection are required. An illustrative embodiment of a method for delivering DNA into maize cells by acceleration is a biolistics α-particle delivery system, which can be used to propel particles coated with DNA through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with corn cells cultured in suspension. Gordon-Kamm *et al.*, describes the basic procedure for coating tungsten particles with DNA (Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990)). The screen disperses the tungsten nucleic acid particles so that they are not delivered to

30

5

10

the recipient cells in large aggregates. A particle delivery system suitable for use with the invention is the helium acceleration PDS-1000/He gun is available from Bio-Rad Laboratories (Bio-Rad, Hercules, California)(Sanford *et al.*, *Technique 3*:3-16 (1991)).

For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein one may obtain up to 1000 or more foci of cells transiently expressing a marker gene. The number of cells in a focus which express the exogenous gene product 48 hours post-bombardment often range from one to ten and average one to three.

In bombardment transformation, one may optimize the pre-bombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for bombardment are important in this technology. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target cells to help alleviate the trauma associated with bombardment and also the nature of the transforming DNA, such as linearized DNA or intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

In another alternative embodiment, plastids can be stably transformed. Methods disclosed for plastid transformation in higher plants include the particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (Svab *et al.*, *Proc. Natl. Acad. Sci.* (*U.S.A.*) 87:8526-8530 (1990); Svab and Maliga, *Proc. Natl. Acad. Sci.* (*U.S.A.*) 90:913-917 (1993); Staub and Maliga, *EMBO J.* 12:601-606 (1993); U.S. Patents 5, 451,513 and 5,545,818).

30

5

10

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small-scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance and helium pressure. One may also minimize the trauma reduction factors by modifying conditions which influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

Agrobacterium-mediated transfer is a widely applicable system for introducing genes into plant cells because the DNA can be introduced into whole plant tissues, thereby bypassing the need for regeneration of an intact plant from a protoplast. The use of Agrobacterium-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art. See, for example the methods described by Fraley et al., Bio/Technology 3:629-635 (1985) and Rogers et al., Methods Enzymol. 153:253-277 (1987). Further, the integration of the Ti-DNA is a relatively precise process resulting in few rearrangements. The region of DNA to be transferred is defined by the border sequences and intervening DNA is usually inserted into the plant genome as described (Spielmann et al., Mol. Gen. Genet. 205:34 (1986)).

Modern Agrobacterium transformation vectors are capable of replication in *E. coli* as well as Agrobacterium, allowing for convenient manipulations as described (Klee et al., In: Plant DNA Infectious Agents, Hohn and Schell (eds.), Springer-Verlag, New York, pp. 179-203 (1985)). Moreover, technological advances in vectors for Agrobacterium-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes and are suitable for present purposes (Rogers et al., Methods Enzymol. 153:253-277 (1987)). In addition, Agrobacterium containing both armed and disarmed Ti genes can be used for the transformations. In those plant strains where Agrobacterium-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer.

30

5

10

A transgenic plant formed using *Agrobacterium* transformation methods typically contains a single gene on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. More preferred is a transgenic plant that is homozygous for the added structural gene; *i.e.*, a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an independent segregant transgenic plant that contains a single added gene, germinating some of the seed produced and analyzing the resulting plants produced for the gene of interest.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating, exogenous genes. Selfing of appropriate progeny can produce plants that are homozygous for both added, exogenous genes that encode a polypeptide of interest. Backcrossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation.

Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation and combinations of these treatments (*See, for example*, Potrykus *et al., Mol. Gen. Genet.* 205:193-200 (1986); Lorz *et al., Mol. Gen. Genet.* 199:178 (1985); Fromm *et al., Nature* 319:791 (1986); Uchimiya *et al., Mol. Gen. Genet.* 204:204 (1986); Marcotte *et al., Nature* 335:454-457 (1988)).

Application of these systems to different plant strains depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from protoplasts are described (Fujimura *et al.*, *Plant Tissue Culture Letters* 2:74 (1985); Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Yamada *et al.*, *Plant Cell Rep.* 4:85 (1986); Abdullah *et al.*, *Biotechnology* 4:1087 (1986)).

To transform plant strains that cannot be successfully regenerated from protoplasts, other ways to introduce DNA into intact cells or tissues can be utilized. For example, regeneration of cereals from immature embryos or explants can be effected as described (Vasil, *Biotechnology* 6:397 (1988)). In addition, "particle gun" or high-velocity microprojectile technology can be utilized (Vasil *et al.*, *Bio/Technology* 10:667 (1992)).

Using the latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al.*, *Nature 328:70* (1987); Klein *et*

30

5

10

al., Proc. Natl. Acad. Sci. (U.S.A.) 85:8502-8505 (1988); McCabe et al., Bio/Technology 6:923 (1988)). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

Other methods of cell transformation can also be used and include but are not limited to introduction of DNA into plants by direct DNA transfer into pollen (Hess *et al.*, *Intern Rev. Cytol. 107*:367 (1987); Luo *et al.*, *Plant Mol Biol. Reporter 6*:165 (1988)), by direct injection of DNA into reproductive organs of a plant (Pena *et al.*, *Nature 325*:274 (1987)), or by direct injection of DNA into the cells of immature embryos followed by the rehydration of desiccated embryos (Neuhaus *et al.*, *Theor. Appl. Genet. 75*:30 (1987)).

The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants are well known in the art (Weissbach and Weissbach, In: *Methods for Plant Molecular Biology*, Academic Press, San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous gene that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens* and obtaining transgenic plants have been published for cotton (U.S. Patent No. 5,004,863; U.S. Patent No. 5,159,135; U.S. Patent No. 5,518,908); soybean (U.S. Patent No. 5,569,834; U.S.

30

5

10

Patent No. 5,416,011; McCabe et. al., Biotechnology 6:923 (1988); Christou et al., Plant Physiol. 87:671-674 (1988)); Brassica (U.S. Patent No. 5,463,174); peanut (Cheng et al., Plant Cell Rep. 15:653-657 (1996), McKently et al., Plant Cell Rep. 14:699-703 (1995)); papaya; and pea (Grant et al., Plant Cell Rep. 15:254-258 (1995)).

Transformation of monocotyledons using electroporation, particle bombardment and Agrobacterium have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier et al., Proc. Natl. Acad. Sci. (USA) 84:5354 (1987)); barley (Wan and Lemaux, Plant Physiol 104:37 (1994)); maize (Rhodes et al., Science 240:204 (1988); Gordon-Kamm et al., Plant Cell 2:603-618 (1990); Fromm et al., Bio/Technology 8:833 (1990); Koziel et al., Bio/Technology 11:194 (1993); Armstrong et al., Crop Science 35:550-557 (1995)); oat (Somers et al., Bio/Technology 10:1589 (1992)); orchard grass (Horn et al., Plant Cell Rep. 7:469 (1988)); rice (Toriyama et al., Theor Appl. Genet. 205:34 (1986); Part et al., Plant Mol. Biol. 32:1135-1148 (1996); Abedinia et al., Aust. J. Plant Physiol. 24:133-141 (1997); Zhang and Wu, Theor. Appl. Genet. 76:835 (1988); Zhang et al., Plant Cell Rep. 7:379 (1988); Battraw and Hall, Plant Sci. 86:191-202 (1992); Christou et al., Bio/Technology 9:957 (1991)); rye (De la Pena et al., Nature 325:274 (1987)); sugarcane (Bower and Birch, Plant J. 2:409 (1992)); tall fescue (Wang et al., Bio/Technology 10:691 (1992)) and wheat (Vasil et al., Bio/Technology 10:667 (1992); U.S. Patent No. 5,631,152).

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte *et al.*, *Nature 335*:454-457 (1988); Marcotte *et al.*, *Plant Cell 1*:523-532 (1989); McCarty *et al.*, *Cell 66*:895-905 (1991); Hattori *et al.*, *Genes Dev. 6*:609-618 (1992); Goff *et al.*, *EMBO J. 9*:2517-2522 (1990)). Transient expression systems may be used to functionally dissect gene constructs (*see generally*, Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)).

Any of the nucleic acid molecules of the invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers, etc. Further, any of the nucleic acid molecules of the invention may be introduced into a plant cell in a manner that allows for overexpression of the protein or fragment thereof encoded by the nucleic acid molecule.

30

5

10



Cosuppression is the reduction in expression levels, usually at the level of RNA, of a particular endogenous gene or gene family by the expression of a homologous sense construct that is capable of transcribing mRNA of the same strandedness as the transcript of the endogenous gene (Napoli *et al.*, *Plant Cell 2*:279-289 (1990); van der Krol *et al.*, *Plant Cell 2*:291-299 (1990)). Cosuppression may result from stable transformation with a single copy nucleic acid molecule that is homologous to a nucleic acid sequence found within the cell (Prolls and Meyer, *Plant J. 2*:465-475 (1992)) or with multiple copies of a nucleic acid molecule that is homologous to a nucleic acid sequence found within the cell (Mittlesten *et al.*, *Mol. Gen. Genet. 244*:325-330 (1994)). Genes, even though different, linked to homologous promoters may result in the cosuppression of the linked genes (Vaucheret, *C.R. Acad. Sci. III 316*:1471-1483 (1993); Flavell, *Proc. Natl. Acad. Sci. (U.S.A.) 91*:3490-3496 (1994)); van Blokland *et al.*, *Plant J. 6*:861-877 (1994); Jorgensen, *Trends Biotechnol. 8*:340-344 (1990); Meins and Kunz, In: *Gene Inactivation and Homologous Recombination in Plants*, Paszkowski (ed.), pp. 335-348, Kluwer Academic, Netherlands (1994)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the cosuppression of an endogenous protein.

Antisense approaches are a way of preventing or reducing gene function by targeting the genetic material (Mol *et al.*, *FEBS Lett.* 268:427-430 (1990)). The objective of the antisense approach is to use a sequence complementary to the target gene to block its expression and create a mutant cell line or organism in which the level of a single chosen protein is selectively reduced or abolished. Antisense techniques have several advantages over other 'reverse genetic' approaches. The site of inactivation and its developmental effect can be manipulated by the choice of promoter for antisense genes or by the timing of external application or microinjection. Antisense can manipulate its specificity by selecting either unique regions of the target gene or regions where it shares homology to other related genes (Hiatt *et al.*, In: *Genetic Engineering*, Setlow (ed.), Vol. 11, New York: Plenum 49-63 (1989)).

The principle of regulation by antisense RNA is that RNA that is complementary to the target mRNA is introduced into cells, resulting in specific RNA:RNA duplexes being formed by base pairing between the antisense substrate and the target mRNA (Green *et al.*, *Annu. Rev*.

10

Biochem. 55:569-597 (1986)). Under one embodiment, the process involves the introduction and expression of an antisense gene sequence. Such a sequence is one in which part or all of the normal gene sequences are placed under a promoter in inverted orientation so that the 'wrong' or complementary strand is transcribed into a noncoding antisense RNA that hybridizes with the target mRNA and interferes with its expression (Takayama and Inouye, Crit. Rev. Biochem. Mol. Biol. 25:155-184 (1990)). An antisense vector is constructed by standard procedures and introduced into cells by transformation, transfection, electroporation, microinjection, infection, etc. The type of transformation and choice of vector will determine whether expression is transient or stable. The promoter used for the antisense gene may influence the level, timing, tissue, specificity, or inducibility of the antisense inhibition.

It is understood that the activity of a protein in a plant cell may be reduced or depressed by growing a transformed plant cell containing a nucleic acid molecule whose non-transcribed strand encodes a protein or fragment thereof.

Posttranscriptional gene silencing (PTGS) can result in virus immunity or gene silencing in plants. PTGS is induced by dsRNA and is mediated by an RNA-dependent RNA polymerase, present in the cytoplasm, that requires a dsRNA template. The dsRNA is formed by hybridization of complementary transgene mRNAs or complementary regions of the same transcript. Duplex formation can be accomplished by using transcripts from one sense gene and one antisense gene co-located in the plant genome, a single transcript that has self-complementarity, or sense and antisense transcripts from genes brought together by crossing. The dsRNA-dependent RNA polymerase makes a complementary strand from the transgene mRNA and RNAse molecules attach to this complementary strand (cRNA). These cRNA-RNAse molecules hybridize to the endogene mRNA and cleave the single-stranded RNA adjacent to the hybrid. The cleaved single-stranded RNAs are further degraded by other host RNAses because one will lack a capped 5' end and the other will lack a poly(A) tail (Waterhouse et al., PNAS 95: 13959-13964 (1998)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the postranscriptional gene silencing of an endogenous transcript.

30

5

10



Antibodies have been expressed in plants (Hiatt *et al.*, *Nature 342:76-78* (1989); Conrad and Fielder, *Plant Mol. Biol. 26:*1023-1030 (1994)). Cytoplasmic expression of a scFv (single-chain Fv antibodies) has been reported to delay infection by artichoke mottled crinkle virus. Transgenic plants that express antibodies directed against endogenous proteins may exhibit a physiological effect (Philips *et al.*, *EMBO J. 16:*4489-4496 (1997); Marion-Poll, *Trends in Plant Science 2:*447-448 (1997)). For example, expressed anti-abscissic antibodies have been reported to result in a general perturbation of seed development (Philips *et al.*, *EMBO J. 16:* 4489-4496 (1997)).

Antibodies that are catalytic may also be expressed in plants (abzymes). The principle behind abzymes is that since antibodies may be raised against many molecules, this recognition ability can be directed toward generating antibodies that bind transition states to force a chemical reaction forward (Persidas, *Nature Biotechnology 15*:1313-1315 (1997); Baca *et al.*, *Ann. Rev. Biophys. Biomol. Struct.* 26:461-493 (1997)). The catalytic abilities of abzymes may be enhanced by site directed mutagenesis. Examples of abzymes are, for example, set forth in U.S. Patent No: 5,658,753; U.S. Patent No. 5,632,990; U.S. Patent No. 5,631,137; U.S. Patent 5,602,015; U.S. Patent No. 5,559,538; U.S. Patent No. 5,576,174; U.S. Patent No. 5,500,358; U.S. Patent 5,318,897; U.S. Patent No. 5,298,409; U.S. Patent No. 5,258,289 and U.S. Patent No. 5,194,585.

It is understood that any of the antibodies of the invention may be expressed in plants and that such expression can result in a physiological effect. It is also understood that any of the expressed antibodies may be catalytic.

The present invention also provides for parts of the plants of the present invention. Plant parts, without limitation, include seed, endosperm, ovule and pollen. In a particularly preferred embodiment of the present invention, the plant part is a seed.

Exemplary Uses

Nucleic acid molecules and fragments thereof of the invention may be employed to obtain other nucleic acid molecules from the same species (nucleic acid molecules from maize may be utilized to obtain other nucleic acid molecules from maize). Such nucleic acid molecules include the nucleic acid molecules that encode the complete coding sequence of a protein and promoters and flanking sequences of such molecules. In addition, such nucleic acid molecules

30

5

10



include nucleic acid molecules that encode for other isozymes or gene family members. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries. Methods for forming such libraries are well known in the art.

Nucleic acid molecules and fragments thereof of the invention may also be employed to obtain nucleic acid homologues. Such homologues include the nucleic acid molecule of other plants or other organisms (e.g., alfalfa, Arabidopsis, barley, Brassica, broccoli, cabbage, citrus, cotton, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm, Phaseolus, etc.) including the nucleic acid molecules that encode, in whole or in part, protein homologues of other plant species or other organisms, sequences of genetic elements, such as promoters and transcriptional regulatory elements. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries obtained from such plant species. Methods for forming such libraries are well known in the art. Such homologue molecules may differ in their nucleotide sequences from those found in one or more of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof because complete complementarity is not needed for stable hybridization. The nucleic acid molecules of the invention therefore also include molecules that, although capable of specifically hybridizing with the nucleic acid molecules may lack "complete complementarity."

Any of a variety of methods may be used to obtain one or more of the above-described nucleic acid molecules (Zamechik et al., Proc. Natl. Acad. Sci. (U.S.A.) 83:4143-4146 (1986); Goodchild et al., Proc. Natl. Acad. Sci. (U.S.A.) 85:5507-5511 (1988); Wickstrom et al., Proc. Natl. Acad. Sci. (U.S.A.) 85:1028-1032 (1988); Holt et al., Molec. Cell. Biol. 8:963-973 (1988); Gerwirtz et al., Science 242:1303-1306 (1988); Anfossi et al., Proc. Natl. Acad. Sci. (U.S.A.) 86:3379-3383 (1989); Becker et al., EMBO J. 8:3685-3691 (1989)). Automated nucleic acid synthesizers may be employed for this purpose. In lieu of such synthesis, the disclosed nucleic acid molecules may be used to define a pair of primers that can be used with the polymerase chain reaction (Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51:263-273 (1986); Erlich et al., European Patent 50,424; European Patent 84,796; European Patent 258,017; European

30

5

10



Patent 237,362; Mullis, European Patent 201,184; Mullis *et al.*, U.S. Patent 4,683,202; Erlich, U.S. Patent 4,582,788; and Saiki *et al.*, U.S. Patent 4,683,194) to amplify and obtain any desired nucleic acid molecule or fragment.

Promoter sequences and other genetic elements, including but not limited to transcriptional regulatory flanking sequences, associated with one or more of the disclosed nucleic acid sequences can also be obtained using the disclosed nucleic acid sequence provided herein. In one embodiment, such sequences are obtained by incubating nucleic acid molecules of the present invention with members of genomic libraries and recovering clones that hybridize to such nucleic acid molecules thereof. In a second embodiment, methods of "chromosome walking," or inverse PCR may be used to obtain such sequences (Frohman *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:8998-9002 (1988); Ohara *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 86*:5673-5677 (1989); Pang *et al.*, *Biotechniques 22*:1046-1048 (1977); Huang *et al.*, *Methods Mol. Biol. 67*:287-294 (1997); Benkel *et al.*, *Genet. Anal. 13*:123-127 (1996); Hartl *et al.*, *Methods Mol. Biol. 58*:293-301 (1996)). The term "chromosome walking" means a process of extending a genetic map by successive hybridization steps.

The nucleic acid molecules of the invention may be used to isolate promoters of cell enhanced, cell specific, tissue enhanced, tissue specific, developmentally or environmentally regulated expression profiles. Isolation and functional analysis of the 5' flanking promoter sequences of these genes from genomic libraries, for example, using genomic screening methods and PCR techniques would result in the isolation of useful promoters and transcriptional regulatory elements. These methods are known to those of skill in the art and have been described (See, for example, Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, (1997), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). Promoters obtained utilizing the nucleic acid molecules of the invention could also be modified to affect their control characteristics. Examples of such modifications would include but are not limited to enhancer sequences. Such genetic elements could be used to enhance gene expression of new and existing traits for crop improvement.

Another subset of the nucleic acid molecules of the invention includes nucleic acid molecules that are markers. The markers can be used in a number of conventional ways in the

30

5

10



field of molecular genetics. Such markers include nucleic acid molecules SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof or fragments of either that can act as markers and other nucleic acid molecules of the present invention that can act as markers.

The genomes of animals and plants naturally undergo spontaneous mutation in the course of their continuing evolution (Gusella, *Ann. Rev. Biochem. 55*:831-854 (1986)). A "polymorphism" is a variation or difference in the sequence of the gene or its flanking regions that arises in some of the members of a species. The variant sequence and the "original" sequence co-exist in the species' population. In some instances, such co-existence is in stable or quasi-stable equilibrium.

A polymorphism is thus said to be "allelic," in that, due to the existence of the polymorphism, some members of a species may have the original sequence (i.e., the original "allele") whereas other members may have the variant sequence (i.e., the variant "allele"). In the simplest case, only one variant sequence may exist and the polymorphism is thus said to be diallelic. In other cases, the species' population may contain multiple alleles and the polymorphism is termed tri-allelic, etc. A single gene may have multiple different unrelated polymorphisms. For example, it may have a di-allelic polymorphism at one site and a multiple polymorphism at another site.

The variation that defines the polymorphism may range from a single nucleotide variation to the insertion or deletion of extended regions within a gene. In some cases, the DNA sequence variations are in regions of the genome that are characterized by short tandem repeats (STRs) that include tandem di- or tri-nucleotide repeated motifs of nucleotides. Polymorphisms characterized by such tandem repeats are referred to as "variable number tandem repeat" ("VNTR") polymorphisms. VNTRs have been used in identity analysis (Weber, U.S. Patent 5,075,217; Armour et al., FEBS Lett. 307:113-115 (1992); Jones et al., Eur. J. Haematol. 39:144-147 (1987); Horn et al., PCT Patent Application WO91/14003; Jeffreys, European Patent Application 370,719; Jeffreys, U.S. Patent 5,175,082; Jeffreys et al., Amer. J. Hum. Genet. 39:11-24 (1986); Jeffreys et al., Nature 316:76-79 (1985); Gray et al., Proc. R. Acad. Soc. Lond. 243:241-253 (1991); Moore et al., Genomics 10:654-660 (1991); Jeffreys et al., Anim. Genet. 18:1-15 (1987); Hillel et al., Anim. Genet. 20:145-155 (1989); Hillel et al., Genet. 124:783-789 (1990)).

5

10



The detection of polymorphic sites in a sample of DNA may be facilitated through the use of nucleic acid amplification methods. Such methods specifically increase the concentration of polynucleotides that span the polymorphic site, or include that site and sequences located either distal or proximal to it. Such amplified molecules can be readily detected by gel electrophoresis or other means.

In an alternative embodiment, such polymorphisms can be detected through the use of a marker nucleic acid molecule that is physically linked to such polymorphism(s). For this purpose, marker nucleic acid molecules comprising a nucleotide sequence of a polynucleotide located within 1 mb of the polymorphism(s) and more preferably within 100kb of the polymorphism(s) and most preferably within 10kb of the polymorphism(s) can be employed.

The identification of a polymorphism can be determined in a variety of ways. By correlating the presence or absence of it in a plant with the presence or absence of a phenotype, it is possible to predict the phenotype of that plant. If a polymorphism creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (e.g., a VNTR polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, individuals that possess a variant sequence can be distinguished from those having the original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed "restriction fragment length polymorphisms" ("RFLPs") (Glassberg, UK Patent Application 2135774; Skolnick *et al.*, *Cytogen. Cell Genet. 32*:58-67 (1982); Botstein *et al.*, *Ann. J. Hum. Genet. 32*:314-331 (1980); Fischer *et al.*, (PCT Application WO90/13668; Uhlen, PCT Application WO90/1369).

Polymorphisms can also be identified by Single Strand Conformation Polymorphism (SSCP) analysis (Elles, *Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases*, Humana Press (1996)); Orita *et al.*, *Genomics* 5:874-879 (1989)). A number of protocols have been described for SSCP including, but not limited to, Lee *et al.*, *Anal. Biochem.* 205:289-293 (1992); Suzuki *et al.*, *Anal. Biochem.* 192:82-84 (1991); Lo *et al.*, *Nucleic Acids Research* 20:1005-1009 (1992); Sarkar *et al.*, *Genomics* 13:441-443 (1992). It is understood that one or more of the nucleic acids of the invention, may be utilized as markers or probes to detect polymorphisms by SSCP analysis.

30

5

10



Polymorphisms may also be found using a DNA fingerprinting technique called amplified fragment length polymorphism (AFLP), which is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA to profile that DNA (Vos *et al.*, *Nucleic Acids Res.* 23:4407-4414 (1995)). This method allows for the specific coamplification of high numbers of restriction fragments, which can be visualized by PCR without knowledge of the nucleic acid sequence. It is understood that one or more of the nucleic acids of the invention, may be utilized as markers or probes to detect polymorphisms by AFLP analysis or for fingerprinting RNA.

Polymorphisms may also be found using random amplified polymorphic DNA (RAPD) (Williams *et al.*, *Nucl. Acids Res. 18*:6531-6535 (1990)) and cleaveable amplified polymorphic sequences (CAPS) (Lyamichev *et al.*, *Science 260*:778-783 (1993)). It is understood that one or more of the nucleic acid molecules of the invention, may be utilized as markers or probes to detect polymorphisms by RAPD or CAPS analysis.

Through genetic mapping, a fine scale linkage map can be developed using DNA markers and, then, a genomic DNA library of large-sized fragments can be screened with molecular markers linked to the desired trait. Molecular markers are advantageous for agronomic traits that are otherwise difficult to tag, such as resistance to pathogens, insects and nematodes, tolerance to abiotic stress, quality parameters and quantitative traits such as high yield potential.

Requirements for marker-assisted selection in a plant breeding program often are: (1) the marker(s) should co-segregate or be closely linked with the desired trait; (2) an efficient means of screening large populations for the molecular marker(s) should be available; and (3) the screening technique should have high reproducibility across laboratories and preferably be economical to use and be user-friendly.

The genetic linkage of marker molecules can be established by a gene mapping model such as, without limitation, the flanking marker model reported by Lander and Botstein, *Genetics 121*:185-199 (1989) and the interval mapping, based on maximum likelihood methods described by Lander and Botstein, *Genetics 121*:185-199 (1989) and implemented in the software package MAPMAKER/QTL (Lincoln and Lander, *Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL*, Whitehead Institute for Biomedical Research, Massachusetts, (1990).

Additional software includes Qgene, Version 2.23 (1996), Department of Plant Breeding and

5

10



Biometry, 266 Emerson Hall, Cornell University, Ithaca, NY). Use of Qgene software is a particularly preferred approach.

A maximum likelihood estimate (MLE) for the presence of a marker is calculated, together with an MLE assuming no QTL effect, to avoid false positives. A \log_{10} of an odds ratio (LOD) is then calculated as: LOD = \log_{10} (MLE for the presence of a QTL/MLE given no linked QTL).

The LOD score essentially indicates how much more likely the data are to have arisen assuming the presence of a QTL than in its absence. The LOD threshold value for avoiding a false positive with a given confidence, say 95%, depends on the number of markers and the length of the genome. Graphs indicating LOD thresholds are set forth in Lander and Botstein, *Genetics 121*:185-199 (1989) and further described by Arús and Moreno-González, *Plant Breeding*, Hayward *et al.*, (eds.) Chapman & Hall, London, pp. 314-331 (1993).

Additional models can be used. Many modifications and alternative approaches to interval mapping have been reported, including the use non-parametric methods (Kruglyak and Lander, Genetics 139:1421-1428 (1995)). Multiple regression methods or models can be also be used, in which the trait is regressed on a large number of markers (Jansen, Biometrics in Plant Breeding, van Oijen and Jansen (eds.), Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp. 116-124 (1994); Weber and Wricke, Advances in Plant Breeding, Blackwell, Berlin, 16 (1994)). Procedures combining interval mapping with regression analysis, whereby the phenotype is regressed onto a single putative QTL at a given marker interval and at the same time onto a number of markers that serve as 'cofactors,' have been reported by Jansen and Stam, Genetics 136:1447-1455 (1994), and Zeng, Genetics 136:1457-1468 (1994). Generally, the use of cofactors reduces the bias and sampling error of the estimated QTL positions (Utz and Melchinger, Biometrics in Plant Breeding, van Oijen and Jansen (eds.) Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp.195-204 (1994), thereby improving the precision and efficiency of QTL mapping (Zeng, Genetics 136:1457-1468 (1994)). These models can be extended to multi-environment experiments to analyze genotype-environment interactions (Jansen et al., Theo. Appl. Genet. 91:33-37 (1995)).

30

5

10



It is understood that one or more of the nucleic acid molecules of the invention may be used as molecular markers. It is also understood that one or more of the protein molecules of the invention may be used as molecular markers.

In accordance with this aspect of the invention, a sample nucleic acid is obtained from plant cells or tissues. Any source of nucleic acid may be used. Preferably, the nucleic acid is genomic DNA. The nucleic acid is subjected to restriction endonuclease digestion. For example, one or more nucleic acid molecule or fragment thereof of the invention can be used as a probe in accordance with the above-described polymorphic methods. The polymorphism obtained in this approach can then be cloned to identify the mutation at the coding region, which alters structure, or regulatory region of the gene, which affects its expression level.

In an aspect of the present invention, one or more of the nucleic molecules of the present invention are used to determine the level (*i.e.*, the concentration of mRNA in a sample, *etc.*) in a plant (preferably maize or soybean) or pattern (*i.e.*, the kinetics of expression, rate of decomposition, stability profile, *etc.*) of the expression of a protein encoded in part or whole by one or more of the nucleic acid molecule of the present invention (collectively, the "Expression Response" of a cell or tissue).

As used herein, the Expression Response manifested by a cell or tissue is said to be "altered" if it differs from the Expression Response of cells or tissues of plants not exhibiting the phenotype. To determine whether a Expression Response is altered, the Expression Response manifested by the cell or tissue of the plant exhibiting the phenotype is compared with that of a similar cell or tissue sample of a plant not exhibiting the phenotype. As will be appreciated, it is not necessary to re-determine the Expression Response of the cell or tissue sample of plants not exhibiting the phenotype each time such a comparison is made; rather, the Expression Response of a particular plant may be compared with previously obtained values of normal plants. As used herein, the phenotype of the organism is any of one or more characteristics of an organism (e.g. disease resistance, pest tolerance, environmental tolerance such as tolerance to abiotic stress, male sterility, quality improvement or yield etc.). A change in genotype or phenotype may be transient or permanent. Also as used herein, a tissue sample is any sample that comprises more than one cell. In a preferred aspect, a tissue sample comprises cells that share a common characteristic (e.g. derived from root, seed, flower, leaf, stem or pollen etc.).

30

5

10

In one aspect of the present invention, an evaluation can be conducted to determine whether a particular mRNA molecule is present. One or more of the nucleic acid molecules of the present invention are utilized to detect the presence or quantity of the mRNA species. Such molecules are then incubated with cell or tissue extracts of a plant under conditions sufficient to permit nucleic acid hybridization. The detection of double-stranded probe-mRNA hybrid molecules is indicative of the presence of the mRNA; the amount of such hybrid formed is proportional to the amount of mRNA. Thus, such probes may be used to ascertain the level and extent of the mRNA production in a plant's cells or tissues. Such nucleic acid hybridization may be conducted under quantitative conditions (thereby providing a numerical value of the amount of the mRNA present). Alternatively, the assay may be conducted as a qualitative assay that indicates either that the mRNA is present, or that its level exceeds a user set, predefined value.

A number of methods can be used to compare the expression response between two or more samples of cells or tissue. These methods include hybridization assays, such as Northerns, RNAse protection assays, and *in situ* hybridization. Alternatively, the methods include PCR-type assays. In a preferred method, the expression response is compared by hybridizing nucleic acids from the two or more samples to an array of nucleic acids. The array contains a plurality of suspected sequences known or suspected of being present in the cells or tissue of the samples.

An advantage of *in situ* hybridization over more conventional techniques for the detection of nucleic acids is that it allows an investigator to determine the precise spatial population (Angerer *et al.*, *Dev. Biol. 101*:477-484 (1984); Angerer *et al.*, *Dev. Biol. 112*:157-166 (1985); Dixon *et al.*, *EMBO J. 10*:1317-1324 (1991)). *In situ* hybridization may be used to measure the steady-state level of RNA accumulation (Hardin *et al.*, *J. Mol. Biol. 202*:417-431 (1989)). A number of protocols have been devised for *in situ* hybridization, each with tissue preparation, hybridization and washing conditions (Meyerowitz, *Plant Mol. Biol. Rep. 5*:242-250 (1987); Cox and Goldberg, In: *Plant Molecular Biology: A Practical Approach*, Shaw (ed.), pp. 1-35, IRL Press, Oxford (1988); Raikhel *et al.*, *In situ RNA hybridization in plant tissues*, In: *Plant Molecular Biology Manual*, vol. B9:1-32, Kluwer Academic Publisher, Dordrecht, Belgium (1989)).

In situ hybridization also allows for the localization of proteins within a tissue or cell (Wilkinson, In Situ Hybridization, Oxford University Press, Oxford (1992); Langdale, In Situ

30

5

10



Hybridization In: The Maize Handbook, Freeling and Walbot (eds.), pp. 165-179, Springer-Verlag, New York (1994)). It is understood that one or more of the molecules of the invention, preferably one or more of the nucleic acid molecules or fragments thereof of the invention or one or more of the antibodies of the invention may be utilized to detect the level or pattern of a protein or mRNA thereof by *in situ* hybridization.

Fluorescent *in situ* hybridization allows the localization of a particular DNA sequence along a chromosome which is useful, among other uses, for gene mapping, following chromosomes in hybrid lines or detecting chromosomes with translocations, transversions or deletions. *In situ* hybridization has been used to identify chromosomes in several plant species (Griffor *et al.*, *Plant Mol. Biol. 17*:101-109 (1991); Gustafson *et al.*, *Proc. Natl. Acad. Sci.* (*U.S.A.*) 87:1899-1902 (1990); Mukai and Gill, *Genome 34*:448-452 (1991); Schwarzacher and Heslop-Harrison, *Genome 34*:317-323 (1991); Wang *et al.*, *Jpn. J. Genet. 66*:313-316 (1991); Parra and Windle, *Nature Genetics 5*:17-21 (1993)). It is understood that the nucleic acid molecules of the invention may be used as probes or markers to localize sequences along a chromosome.

Another method to localize the expression of a molecule is tissue printing. Tissue printing provides a way to screen, at the same time on the same membrane many tissue sections from different plants or different developmental stages (Yomo and Taylor, *Planta 112*:35-43 (1973); Harris and Chrispeels, *Plant Physiol*. 56:292-299 (1975); Cassab and Varner, *J. Cell. Biol*. 105:2581-2588 (1987); Spruce et al., *Phytochemistry 26*:2901-2903 (1987); Barres et al., *Neuron 5*:527-544 (1990); Reid and Pont-Lezica, *Tissue Printing: Tools for the Study of Anatomy, Histochemistry and Gene Expression*, Academic Press, New York, New York (1992); Reid et al., *Plant Physiol*. 93:160-165 (1990); Ye et al., *Plant J. 1*:175-183 (1991)).

It is understood that one or more of the molecules of the invention, preferably one or more of the nucleic acid molecules of the present invention or one or more of the antibodies of the invention may be utilized to detect the presence or quantity of a protein or fragment of the invention by tissue printing.

Further it is also understood that any of the nucleic acid molecules of the invention may be used as marker nucleic acids and or probes in connection with methods that require probes or marker nucleic acids. As used herein, a probe is an agent that is utilized to determine an attribute

5

10



or feature (e.g. presence or absence, location, correlation, etc.) of a molecule, cell, tissue or plant. As used herein, a marker nucleic acid is a nucleic acid molecule that is utilized to determine an attribute or feature (e.g., presence or absence, location, correlation, etc.) or a molecule, cell, tissue or plant.

A microarray-based method for high-throughput monitoring of gene expression may be utilized to measure expression response Schena *et al.*, *Science* 270:467-470 (1995); http://cmgm.stanford.edu/pbrown/array.html; Shalon, Ph.D. Thesis, Stanford University (1996). This approach is based on using arrays of DNA targets (e.g. cDNA inserts, colonies, or polymerase chain reaction products) for hybridization to a "complex probe" prepared with RNA extracted from a given cell line or tissue. The probe may be produced by reverse transcription of mRNA or total RNA and labeled with radioactive or fluorescent labeling. The probe is complex in that it contains many different sequences in various amounts, corresponding to the numbers of copies of the original mRNA species extracted from the sample.

The initial RNA source will typically be derived from a physiological source. The physiological source may be derived from a variety of eukaryotic sources, with physiological sources of interest including sources derived from single celled organisms such as yeast and multicellular organisms, including plants and animals, particularly plants, where the physiological sources from multicellular organisms may be derived from particular organs or tissues of the multicellular organism, or from isolated cells derived therefrom. The physiological sources may be derived from multicellular organisms at different developmental stages (e.g., 10-day-old seedlings), grown under different environmental conditions (e.g., drought-stressed plants) or treated with chemicals.

In obtaining the sample of RNAs to be analyzed from the physiological source from which it is derived, the physiological source may be subjected to a number of different processing steps, where such processing steps might include tissue homogenation, cell isolation and cytoplasmic extraction, nucleic acid extraction and the like, where such processing steps are known to the those of skill in the art. Methods of isolating RNA from cells, tissues, organs or whole organisms are known to those of skill in the art and are described in Maniatis et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Press) (1989).

30

5

10



The DNA may be placed on nylon or glass "microarrays" regularly arranged with a spot spacing of 1 mm or less. Expression levels can be measured for hundreds or thousands of genes, by using less than 2 micrograms of polyA+ RNA and determining the relative mRNA abundances down to one in ten thousand or less (Granjeaud *et. al., BioEssays* 21:781-790 (1999)).

In addition to arrays of cDNA clones or inserts, arrays of oligonucleotides are also used to study differential gene expression. In an oligonucleotide array, the genes of interest are represented by a series of approximately 20 nucleotide oligomers that are unique to each gene. Labeled mRNA is prepared and hybridization signals are detected from specific sets of oligos that represent different genes supplemented by a set of control oligonucleotides. Potential advantages of the oligonucleotide array include enhanced specificity and sensitivity through the parallel analysis of "perfect match" oligos and "mismatch" oligos for each gene. The hybridization conditions can be adjusted to distinguish a perfect heteroduplex from a single base mismatch, thus allowing subtraction of nonspecific hybridization signals from specific hybridization signals. A disadvantage of oligonucleotide arrays relative to cDNA arrays is the limitation of the technology to genes of known sequence (Granjeaud *et. al.*, *BioEssays* 21:781-790 (1991); Carulli *et al.*, *Journal of Cellular Biochemistry Supplements* 30/31:286-296 (1998)).

These techniques have been successfully used to characterize patterns of gene expression associated with, for example, various important physiological changes in yeast, including the mitotic cell cycle, the heat shock response, and comparison between mating types. Once a set of comparable expression profiles is obtained, e.g. for cells at different time points or at different cellular states, a clustering algorithm generally is used to group sets of genes which share similar expression patterns. The clusters obtained can then be analyzed in the light of available functional annotations, often leading to associations of poorly characterized genes with genes whose function and regulation are better understood.

Regulatory networks that control gene expression can be characterized using microarray technology (DeRisi *et al.*, *Science* 278: 680-686 (1997); Winzler *et al. Science* 28: 1194-1197 (1998); Cho *et al. Mol Cell* 2: 65-73 (1998); Spellman *et al. Mol Biol Cell* 95: 14863-14868 (1998). For example, it is has been reported that both cDNA and oligonucleotide arrays have been used to monitor gene expression in synchronized cell cultures. Analysis of the

30

5

10



corresponding temporal patterns of gene expression resulted in the identification of over 400 cell cycle-regulated genes. In order to identify possible common regulatory mechanisms accounting for co-expression, consensus motifs in putative regulatory sequences upstream of the corresponding ORFs were examined. This resulted in the identification of several new potential binding sites for known factors or complexes involved in the coordinated transcription of genes during specific phases of the cell cycle (Thieffry, D. *BioEssays 21*: 895-899 (1999)).

The microarray approach may be used with polypeptide targets (U.S. Patent No. 5,445,934; U.S. Patent No. 5,143,854; U.S. Patent No. 5,079,600; U.S. Patent No. 4,923,901) synthesized on a substrate (microarray) and these polypeptides can be screened with either (Fodor *et al.*, *Science 251*:767-773 (1991)). It is understood that one or more of the nucleic acid molecules or protein or fragments thereof of the invention may be utilized in a microarray-based method.

In a preferred embodiment of the present invention microarrays may be prepared that comprise nucleic acid molecules where preferably at least 10%, preferably at least 25%, more preferably at least 50% and even more preferably at least 75%, 80%, 85%, 90% or 95% of the nucleic acid molecules located on that array are selected from the group of nucleic acid molecules that specifically hybridize to one or more nucleic acid molecule having a nucleic acid sequence selected from the group of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complement thereof or fragments of either.

In another preferred embodiment of the present invention microarrays may be prepared that comprise nucleic acid molecules where preferably at least 10%, preferably at least 25%, more preferably at least 50% and even more preferably at least 75%, 80%, 85%, 90% or 95% of the nucleic acid molecules located on that array are selected from the group of nucleic acid molecules having a nucleic acid sequence selected from the group of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof.

In an even more preferred embodiment of the present invention, the microarray comprises a nucleic acid molecule and/or collection of nucleic acid molecules of the present invention where the nucleic acid molecule and/or collection of nucleic acid molecules are capable of determining or predicting a component or attribute of a biochemical process or activity where the process or activity is preferably selected from photosynthetic activity,

30

5

10



carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, and lipid metabolism, and more preferably selected from the group consisting of biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis and gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate synthesis/degradation, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine metabolism, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid synthesis metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, alpha-oxidation lipid metabolism, and fatty acid metabolism, and even more preferably selected from the group consisting of: glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

In an even more preferred embodiment of the present invention, the microarray comprises a nucleic acid molecule and/or collection of nucleic acid molecules of the present invention where the nucleic acid molecule and/or collection of nucleic acid molecules are capable of detecting or predicting a component or attribute of at least two, more preferable at least three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty one, twenty two, twenty three, twenty four, twenty five, twenty six, twenty seven, twenty eight, twenty nine, thirty, thirty one, thirty two, thirty three, thirty four, thirty five, thirty six, thirty seven, thirty eight, thirty nine, forty, forty one, forty two, forty three, forty four, forty five or forty six biochemical processes or activities where the biochemical processes or activities are selected from the following: photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, lipid metabolism, biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis and gluconeogenesis

5

10



metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate synthesis/degradation, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, â-oxidation lipid metabolism, fatty acid metabolism, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

Site directed mutagenesis may be utilized to modify nucleic acid sequences, particularly as it is a technique that allows one or more of the amino acids encoded by a nucleic acid molecule to be altered (e.g., a threonine to be replaced by a methionine) (Wells et al., Gene 34:315-323 (1985); Gilliam et al., Gene 12:129-137 (1980); Zoller and Smith, Methods Enzymol. 100:468-500 (1983); Dalbadie-McFarland et al., Proc. Natl. Acad. Sci. (U.S.A.) 79:6409-6413 (1982); Scharf et al., Science 233:1076-1078 (1986); Higuchi et al., Nucleic Acids Res. 16:7351-7367 (1988); U.S. Patent 5,811,238, European Patent 0 385 962; European Patent 0 359 472; and PCT Patent Application WO 93/07278; Lanz et al., J. Biol. Chem. 266:9971-9976 (1991); Kovgan and Zhdanov, Biotekhnologiya 5:148-154, No. 207160n, Chemical Abstracts 110:225 (1989); Ge et al., Proc. Natl. Acad. Sci. (U.S.A.) 86:4037-4041 (1989); Zhu et al., J. Biol. Chem. 271:18494-18498 (1996); Chu et al., Biochemistry 33:6150-6157 (1994); Small et al., EMBO J. 11:1291-1296 (1992); Cho et al., Mol. Biotechnol. 8:13-16 (1997); Kita et al., J. Biol. Chem. 271:26529-26535 (1996); Jin et al., Mol. Microbiol. 7:555-562 (1993); Hatfield and Vierstra, J. Biol. Chem. 267:14799-14803 (1992); Zhao et al., Biochemistry 31:5093-5099 (1992)).

Any of the nucleic acid molecules of the invention may either be modified by site directed mutagenesis or used as, for example, nucleic acid molecules that are used to target other nucleic acid molecules for modification.

30

5

10



It is understood that mutants with more than one altered nucleotide can be constructed using techniques that practitioners are familiar with, such as isolating restriction fragments and ligating such fragments into an expression vector (*see*, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989)).

Two steps may be employed to characterize DNA-protein interactions. The first is to identify sequence fragments that interact with DNA-binding proteins, to titrate binding activity, to determine the specificity of binding and to determine whether a given DNA-binding activity can interact with related DNA sequences (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989)). Electrophoretic mobility-shift assay is a widely used assay. The assay provides a rapid and sensitive method for detecting DNA-binding proteins based on the observation that the mobility of a DNA fragment through a nondenaturing, low-ionic strength polyacrylamide gel is retarded upon association with a DNA-binding protein (Fried and Crother, *Nucleic Acids Res.* 9:6505-6525 (1981)). When one or more specific binding activities have been identified, the exact sequence of the DNA bound by the protein may be determined.

Several procedures for characterizing protein/DNA-binding sites are used (Maxam and Gilbert, *Methods Enzymol.* 65:499-560 (1980); Wissman and Hillen, *Methods Enzymol.* 208:365-379 (1991); Galas and Schmitz, *Nucleic Acids Res.* 5:3157-3170 (1978); Sigman *et al.*, *Methods Enzymol.* 208:414-433 (1991); Dixon *et al.*, *Methods Enzymol.* 208:414-433 (1991)). It is understood that one or more of the nucleic acid molecules of the invention may be utilized to identify a protein or fragment thereof that specifically binds to a nucleic acid molecule of the invention. It is also understood that one or more of the protein molecules or fragments thereof of the invention may be utilized to identify a nucleic acid molecule that specifically binds to it.

A two-hybrid system is based on the fact that proteins, such as transcription factors that interact (physically) with one another carry out many cellular functions. Two-hybrid systems have been used to probe the function of new proteins (Chien *et al.*, *Proc. Natl. Acad. Sci.* (*U.S.A.*) 88:9578-9582 (1991); Durfee *et al.*, *Genes Dev.* 7:555-569 (1993); Choi *et al.*, *Cell* 78:499-512 (1994); Kranz *et al.*, *Genes Dev.* 8:313-327 (1994)).

Interaction mating techniques have facilitated a number of two-hybrid studies of proteinprotein interaction. Interaction mating has been used to examine interactions between small sets

30

5

10



of tens of proteins (Finley and Brent, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:12098-12984 (1994)), larger sets of hundreds of proteins (Bendixen *et al.*, *Nucl. Acids Res.* 22:1778-1779 (1994)) and to comprehensively map proteins encoded by a small genome (Bartel *et al.*, *Nature Genetics* 12:72-77 (1996)). This technique utilizes proteins fused to the DNA-binding domain and proteins fused to the activation domain. They are expressed in two different haploid yeast strains of opposite mating type and the strains are mated to determine if the two proteins interact. Mating occurs when haploid yeast strains come into contact and result in the fusion of the two haploids into a diploid yeast strain. An interaction can be determined by the activation of a two-hybrid reporter gene in the diploid strain.

It is understood that the protein-protein interactions of protein or fragments thereof of the invention may be investigated using the two-hybrid system and that any of the nucleic acid molecules of the invention that encode such proteins or fragments thereof may be used to transform yeast in the two-hybrid system.

Computer Readable Media

The nucleotide sequence provided in SEQ ID NO: 1 through SEQ ID NO: 463,173 or fragment thereof, or complement thereof, or a nucleotide sequence at least 90% identical, preferably 95%, identical even more preferably 99% or 100% identical to the sequence provided in SEQ ID NO: 1 through SEQ ID NO: 463,173 or fragment thereof, or complement thereof, can be "provided" in a variety of mediums to facilitate use. Such a medium can also provide a subset thereof in a form that allows a skilled artisan to examine the sequences.

In a preferred embodiment of the present invention computer readable media may be prepared that comprise nucleic acid sequences where preferably at least 10%, preferably at least 25%, more preferably at least 50% and even more preferably at least 75%, 80%, 85%, 90% or 95% of the nucleic acid sequences are selected from the group of nucleic acid molecules that specifically hybridize to one or more nucleic acid molecule having a nucleic acid sequence selected from the group of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complement thereof or fragments of either.

In another preferred embodiment of the present invention computer readable media may be prepared that comprise nucleic acid sequences where preferably at least 10%, preferably at least 25%, more preferably at least 50% and even more preferably at least 75%, 80%, 85%, 90%

30

5

10

or 95% of the nucleic acid sequences are selected from the group of nucleic acid molecules having a nucleic acid sequence selected from the group of SEQ ID NO: 1 through SEQ ID NO: 463,173 or complements thereof.

In a more preferred embodiment of the present invention, the computer readable media comprises a nucleic acid sequence and/or collection of nucleic acid sequences of the present invention associated with a biochemical process or activity where the process or activity is preferably selected from photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, and lipid metabolism, and more preferably selected from the group consisting of biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis and gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate synthesis/degradation, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine metabolism, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, α-oxidation lipid metabolism, and fatty acid metabolism, and even more preferably selected from the group consisting of: glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

In an even more preferred embodiment of the present invention, the computer readable media comprises a nucleic acid sequence and/or collection of nucleic acid sequences of the present invention where the nucleic acid sequence and/or collection of nucleic acid sequences are associated with a component or attribute of at least two, more preferable at least three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty one, twenty two, twenty three, twenty four, twenty five, twenty six, twenty seven, twenty eight, twenty nine, thirty, thirty one, thirty two, thirty three,

30

5

10

thirty four, thirty five, thirty six, thirty seven, thirty eight, thirty nine, forty, forty one, forty two, forty three, forty four, forty five or forty six biochemical processes or activities where the biochemical processes or activities are selected from the following: photosynthetic activity, carbohydrate metabolism, amino acid synthesis or degradation, plant hormone or other regulatory molecules, phenolic metabolism, lipid metabolism, biosynthesis of tetrapyrroles, phytochrome metabolism, carbon assimilation, glycolysis and gluconeogenesis metabolism, sucrose metabolism, starch metabolism, phosphogluconate metabolism, galactomannan metabolism, raffinose metabolism, complex carbohydrate synthesis/degradation, phytic acid metabolism, methionine biosynthesis, methionine degradation, lysine metabolism, arginine metabolism, proline metabolism, glutamate/glutamine, aspartate/asparagine metabolism, cytokinin metabolism, gibberellin metabolism, ethylene metabolism, jasmonic acid synthesis metabolism, transcription factors, R-genes, plant proteases, protein kinases, antifungal proteins, nitrogen transporters, sugar transporters, shikimate metabolism, isoflavone metabolism, phenylpropanoid metabolism, isoprenoid metabolism, β-oxidation lipid metabolism, fatty acid metabolism, glycolysis metabolism, gluconeogenesis metabolism, sucrose metabolism, sucrose catabolism, reductive pentose phosphate cycle, regulation of C3 photosynthesis, C4 pathway carbon assimilation, enzymes involved in the C4 pathway, carotenoid metabolism, tocopherol metabolism, phytosterol metabolism, brassinoid metabolism, and proline metabolism.

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc, storage medium and magnetic tape: optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate media comprising the

30

5

10

nucleotide sequence information of the present invention. A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing one or more of nucleotide sequences of the present invention, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul *et al.*, *J. Mol. Biol. 215*:403-410 (1990)) and BLAZE (Brutlag *et al.*, *Comp. Chem. 17*:203-207 (1993)) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) within the genome that contain homology to ORFs or proteins from other organisms. Such ORFs are protein-encoding fragments within the sequences of the present invention and are useful in producing commercially important proteins such as enzymes used in amino acid biosynthesis, metabolism, transcription, translation, RNA processing, nucleic acid and a protein degradation, protein modification and DNA replication, restriction, modification, recombination and repair.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify commercially important fragments of the nucleic acid molecule of the present invention. As used herein, "a computer-based system" refers to the hardware means, software means and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a

5

10

central processing unit (CPU), input means, output means and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention.

As indicated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory that can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention. As used herein, "search means" refers to one or more programs that are implemented on the computerbased system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequence of the present invention that match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are available can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, MacPattern (EMBL), BLASTIN and BLASTIX (NCBIA). One of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that during searches for commercially important fragments of the nucleic acid molecules of the present invention, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequences the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymatic active sites and signal sequences. Nucleic acid target motifs

30

5

10



include, but are not limited to, promoter sequences, *cis* elements, hairpin structures and inducible expression elements (protein binding sequences).

Thus, the present invention further provides an input means for receiving a target sequence, a data storage means for storing the target sequences of the present invention sequence identified using a search means as described above and an output means for outputting the identified homologous sequences. A variety of structural formats for the input and output means can be used to input and output information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the sequence of the present invention by varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments sequence of the present invention. For example, implementing software that implements the BLAST and BLAZE algorithms (Altschul *et al.*, *J. Mol. Biol. 215*:403-410 (1990)) can be used to identify open frames within the nucleic acid molecules of the present invention. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration and are not intended to be limiting of the present invention, unless specified.

Example 1

This example illustrates the generation of libraries from cDNA prepared from a variety of Arabidopsis thaliana, Columbia ecotype, Landsberg ecotype, Nossen ecotype, Glycine max, Zea mays L., Gossympium hirsutum, Sorghum bicolor, Oryza sativa L (japonica type), Oryza sativa L (japonica type), cv. Nipponbare, Zea mays L. ssp mexicana and Triticum aestivum tissue. A subset of Arabidopsis libraries is used as an example.

Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. Tissue is harvested as follows:

30

5

10



- (a) For leaf tissue-based cDNA, leaf blades are cut with sharp scissors at seven weeks after planting;
- (b) For root tissue-based cDNA, roots of seven-week old plants are rinsed intensively with tap water to wash away dirt, and briefly blotted by paper towel to take away free water;
- (c) For stem tissue-based cDNA, stems are collected seven to eight weeks after planting by cutting the stems from the base and cutting the top of the plant to remove the floral tissue;
- (d) For flower bud tissue-based cDNA, green and unopened flower buds are harvested about seven weeks after planting;
- (e) For open flower tissue-based cDNA, completely opened flowers with all parts of floral structure observable, but no siliques are appearing, and are harvested about seven weeks after planting;
- (f) For immature seed tissue-based cDNA, seeds are harvested at approximately 7-8 weeks of age. The seeds range in maturity from the smallest seeds that could be dissected from siliques to just before starting to turn yellow in color.

All tissue is immediately frozen in liquid nitrogen and stored at -80°C until total RNA extraction. The stored RNA is purified using Trizol reagent from Life Technologies (Gibco BRL, Life Technologies, Gaithersburg, Maryland U.S.A.), essentially as recommended by the manufacturer. Poly A+ RNA (mRNA) is purified using magnetic oligo dT beads essentially as recommended by the manufacturer (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.).

Construction of plant cDNA libraries is well known in the art and a number of cloning strategies exist. A number of cDNA library construction kits are commercially available. The Superscript™ Plasmid System for cDNA synthesis and Plasmid Cloning (Gibco BRL, Life Technologies, Gaithersburg, Maryland U.S.A.) is used, following the conditions suggested by the manufacturer.

The cDNA libraries are plated on LB agar containing the appropriate antibiotics for selection and incubated at 37° for a sufficient time to allow the growth of individual colonies.

30

5

10



Single selective media colonies are individually placed in each well of a 96-well microtiter plates containing LB liquid including the selective antibiotics. The plates are incubated overnight at approximately 37°C with gentle shaking to promote growth of the cultures. The plasmid DNA is isolated from each clone using Qiaprep plasmid isolation kits, using the conditions recommended by the manufacturer (Qiagen Inc., Santa Clara, California U.S.A.).

The template plasmid DNA clones are used for subsequent sequencing. For sequencing the cDNA libraries, a commercially available sequencing kit, such as the ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS, is used under the conditions recommended by the manufacturer (PE Applied Biosystems, Foster City, CA). The cDNAs of the present invention are generated by sequencing initiated from the 5' end or 3' end of each cDNA clone. Entire inserts or only part of the inserts (ESTs or expressed sequenced tags) are sequenced.

A number of sequencing techniques are known in the art, including fluorescence-based sequencing methodologies. These methods have the detection, automation and instrumentation capability necessary for the analysis of large volumes of sequence data. Currently, the 377 DNA Sequencer (Perkin-Elmer Corp., Applied Biosystems Div., Foster City, CA) allows the most rapid electrophoresis and data collection. With these types of automated systems, fluorescent dye-labeled sequence reaction products are detected and data entered directly into the computer, producing a chromatogram that is subsequently viewed, stored, and analyzed using the corresponding software programs. These methods are known to those of skill in the art and have been described and reviewed (Birren et al., Genome Analysis: Analyzing DNA,1, Cold Spring Harbor, New York, the entirety of which is herein incorporated by reference).

The generated ESTs (including any full-length cDNA inserts or complete coding sequences) are combined with ESTs and full length cDNA sequences in public databases such as GenBank. Duplicate sequences are removed; and, duplicate sequence identification numbers are replaced. The combined dataset is then clustered and assembled using Pangea Systems tool identified as CAT v.3.2. First, the EST sequences are screened and filtered, e.g. high frequency words are masked to prevent spurious clustering; sequence common to known contaminants such as cloning bacteria are masked; high frequency repeated sequences and simple sequences are masked; unmasked sequences of less than 100 bp are eliminated. The thus-screened and filtered

5

10

ESTs are combined and subjected to a word-based clustering algorithm which calculates sequence pair distances based on word frequencies and uses a single linkage method to group like sequences into clusters of more than one sequence, as appropriate. Clustered sequence files are assembled individually using an iterative method based on PHRAP/CRAW/MAP providing one or more self-consistent consensus sequences and inconsistent singleton sequences. The assembled clustered sequence files are checked for completeness and parsed to create data representing each consensus contiguous sequence (contig), the initial EST sequences, and the relative position of each EST in a respective contig. The sequence of the 5' most clone is identified from each contig. The initial sequences that are not included in a contig are separated out. A FASTA file is created consisting of 463,173 sequences comprising the sequence of each contig and all original sequences which were not included in a contig. The EST contigs and original sequences which are not included in a contig are presented in Table 1 comprising SEQ ID No: 1 through SEQ ID NO: 463,173.

Example 2

The GenBank database is searched with BLASTN, version 2.0 (BLASTN takes a nucleotide sequence (the query sequence) and its reverse complement and searches them against a nucleotide sequence database) and BLASTX version 2.0 (BLASTX takes a nucleotide sequence, translates it in three forward reading frames and three reverse complement reading frames, and then compares the six translations against a protein sequence database) using default values with the cDNAs as queries. cDNA nucleic acid molecules that pass the E value threshold of $10e^{-8}$ for the following enzymes are classified. Results from these searches are set forth in Table 1.

References

Each reference mentioned in this specification is incorporated by reference in its entirety. In addition, these references, as well as each of those cited can be relied upon to make and use aspects of the invention.



Glycine max

5'-most EST

Method

```
Seq. No.
                   0 1.R1040
Contig ID
5'-most EST
                   CPR2866_700213140_FL
Method
                   BLASTX
NCBI GI
                   g4262236
BLAST score
                   405
E value
                   4.0e-64
Match length
                   214
% identity
NCBI Description
                   (AC006200) putative ribose 5-phosphate isomerase
                   [Arabidopsis thaliana]
Seq. No.
                   1 1.R1040
Contig ID
5'-most EST
                   sat701014721.h1
Method
                   BLASTX
NCBI GI
                   g2980802
BLAST score
                   400
E value
                   2.0e-38
Match length
                   189
% identity
                   46
NCBI Description
                   (AL022197) transcriptional activator CBF1 - like protein
                   [Arabidopsis thaliana]
Seq. No.
                   1 2.R1040
Contig ID
5'-most EST
                  CPR6130_700750838_FL
Method
                  BLASTX
NCBI GI
                   g3660548
BLAST score
                   207
E value
                   5.0e-16
Match length
                  145
% identity
                  35
NCBI Description
                  (AB013815) DREB1A [Arabidopsis thaliana]
Seq. No.
Contig ID
                  1 3.R1040
5'-most EST
                  uC-gmropic021g01b1
Seq. No.
Contig ID
                  2 1.R1040
                  rlr700897446.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q2980802
BLAST score
                  326
E value
                  8.0e-30
Match length
                  212
% identity
                  38
                  (AL022197) transcriptional activator CBF1 - like protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
Contig ID
                  3 1.R1040
```

CPR6164 700155735 FL

BLASTX

```
NCBI GI
                   g4105772
                   804
BLAST score
                   6.0e-86
E value
Match length
                  183
                   53
% identity
                  (AF049917) PGP9B [Petunia x hybrida]
NCBI Description
Seq. No.
                   4 1.R1040
Contig ID
                  ssr700556639.h1
5'-most EST
                  BLASTX
Method
                   g2655008
NCBI GI
BLAST score
                   867
                   5.0e-93
E value
Match length
                   374
                   45
% identity
                   (AF017144) (1-4)-beta-mannan endohydrolase [Lycopersicon
NCBI Description
                   esculentum]
                   8
Seq. No.
                   4 2.R1040
Contig ID
5'-most EST
                   jC-gmle01810053b07a1
                   BLASTX
Method
                   q2642640
NCBI GI
BLAST score
                   1554
                   1.0e-173
E value
Match length
                   339
% identity
                   85
                   (AF032473) ADP-glucose pyrophosphorylase large subunit 2
NCBI Description
                   [Citrullus lanatus]
                   9
Seq. No.
                   4 3.R1040
Contig ID
5'-most EST
                   jC-gmf102220050f03d1
Seq. No.
                   10
                   4 4.R1040
Contig ID
                   uC-gmflminsoy075b06b1
5'-most EST
                   BLASTX
Method
                   g2655008
NCBI GI
BLAST score
                   313
                   1.0e-28
E value
Match length
                   101
% identity
                   54
                   (AF017144) (1-4)-beta-mannan endohydrolase [Lycopersicon
NCBI Description
                   esculentum]
                   11
Seq. No.
                   5 1.R1040
Contig ID
                   LTB3106-062-Q1-K1-E6
5'-most EST
                   BLASTN
Method
NCBI GI
                   g1663540
BLAST score
                   277
E value
                   1.0e-154
Match length
                   461
```

NCBI Description Glycine max clone RLG3 disease resistance protein homolog

90

% identity



gene, partial cds

```
      Seq. No.
      12

      Contig ID
      6_1.R1040

      5'-most EST
      ASG32442DA-001-B5_FL

      Method
      BLASTX

      NCBI GI
      g1709127

      BLAST score
      2023

      E value
      0.0e+00

      Match length
      409
```

93

NCBI Description GLYCOGEN SYNTHASE KINASE-3 HOMOLOG MSK-1

>gi_481020_pir__S37644 protein kinase MSK-1 (EC 2.7.1.-) - alfalfa >gi_313144_emb_CAA48474_ (X68411) protein kinase

[Medicago sativa]

Seq. No. 13 Contig ID 6 2

% identity

6 2.R1040

5'-most EST jC-gmro02910008h06a1

Method BLASTN
NCBI GI g1431621
BLAST score 178
E value 4.0e-95
Match length 527
% identity 84

NCBI Description T.repens mRNA for protein kinase

Seq. No. 14

Contig ID 6 4.R1040

5'-most EST ncj700980733.h1

Method BLASTN
NCBI GI g313143
BLAST score 109
E value 5.0e-54
Match length 485

Match length 485 % identity 86

NCBI Description A.medicago MSK-1 mRNA for protein kinase

Seq. No. 15

Contig ID 6 5.R1040

5'-most EST $j\overline{C}$ -gmfl02220050g09a1

Method BLASTN
NCBI GI g313143
BLAST score 109
E value 3.0e-54
Match length 153
% identity 93

NCBI Description A.medicago MSK-1 mRNA for protein kinase

Seq. No. 16

Contig ID 6 6.R1040

5'-most EST uC-gmrominsoy213e03b1

Method BLASTN
NCBI GI g313143
BLAST score 216
E value 1.0e-118
Match length 359



NCBI Description A.medicago MSK-1 mRNA for protein kinase

Seq. No. 17

Contig ID

6 9.R1040

5'-most EST

% identity

leu701150045.hl

Method

BLASTN

NCBI GI

q313143

BLAST score

169

E value

5.0e-90

Match length

224

% identity

94

NCBI Description A.medicago MSK-1 mRNA for protein kinase

Seq. No.

18

Contig ID

7 1.R1040

5'-most EST

LIB3106-035-Q1-K1-G4

Method NCBI GI BLASTX g1345977

BLAST score

1965

E value

Match length

0.0e + 00383

% identity

94

NCBI Description

OMEGA-6 FATTY ACID DESATURASE, ENDOPLASMIC RETICULUM

ISOZYME 2 >gi_904154 (L43921) microsomal omega-6 desaturase

[Glycine max]

Seq. No.

19

Contig ID

7 2.R1040

5'-most EST

LIB3053-001-Q1-B1-H3

Method

BLASTX

NCBI GI

g1345977

BLAST score

1939

E value

0.0e + 00

383

Match length

93

% identity NCBI Description

OMEGA-6 FATTY ACID DESATURASE, ENDOPLASMIC RETICULUM

ISOZYME 2 >gi 904154 (L43921) microsomal omega-6 desaturase

[Glycine max]

Seq. No.

20

Contig ID

7 3.R1040

5'-most EST

zhf700958467.h1

Method

BLASTX

NCBI GI

g2673906 525

BLAST score E value

1.0e-53

139

Match length % identity

NCBI Description

(AC002561) putative DNA polymerase delta small subunit

[Arabidopsis thaliana]

Seq. No.

Contig ID

7 4.R1040

5'-most EST

uC-gmrominsoy162c08b1

Method

BLASTX

NCBI GI

g2673906



BLAST score 314 E value 3.0e-35 Match length 97 % identity 84

NCBI Description (AC002561) putative DNA polymerase delta small subunit

[Arabidopsis thaliana]

Seq. No. 22

Contig ID 7_5.R1040

5'-most EST kll701211444.h1

Method BLASTN
NCBI GI g904153
BLAST score 97
E value 3.0e-47
Match length 158
% identity 93

NCBI Description Glycine max FAD2-2 microsomal omega-6 desaturase mRNA,

complete cds

Seq. No. 23

Contig ID 8_1.R1040

5'-most EST uC-gmrominsoy222b04b1

Method BLASTX
NCBI GI g3452497
BLAST score 2296
E value 0.0e+00
Match length 566
% identity 79

NCBI Description (Y17796) ketol-acid reductoisomerase [Pisum sativum]

Seq. No. 24

Contig ID 8 2.R1040

5'-most EST jC-gmro02910060h03a1

Method BLASTN
NCBI GI g4210834
BLAST score 71
E value 2.0e-31
Match length 115
% identity 90

NCBI Description Pisum sativum mRNA for ketol-acid reductoisomerase

Seq. No. 25

Contig ID 8 3.R1040

5'-most EST LIB3049-016-Q1-E1-E11

Method BLASTN
NCBI GI g4210834
BLAST score 213
E value 1.0e-116
Match length 569
% identity 84

NCBI Description Pisum sativum mRNA for ketol-acid reductoisomerase

Seq. No. 26

Contig ID 8_5.R1040

5'-most EST jC-gmro02910073g04d1

Method BLASTN NCBI GI g4210834



BLAST score 64
E value 2.0e-27
Match length 116
% identity 89

NCBI Description Pisum sativum mRNA for ketol-acid reductoisomerase

Seq. No. 8 6.R1040 Contig ID 5'-most EST g4304571 Method BLASTN NCBI GI q4210834 BLAST score 157 7.0e-83 E value Match length 376 % identity 88

NCBI Description Pisum sativum mRNA for ketol-acid reductoisomerase

Seq. No. 28

Contig ID 9 1.R1040

5'-most EST leu701146904.h1

Method BLASTX
NCBI GI g2842490
BLAST score 791
E value 1.0e-106
Match length 337
% identity 62

NCBI Description (AL021749) heat-shock protein [Arabidopsis thaliana]

Seq. No. 29

Contig ID 9 2.R1040

5'-most EST LIB3049-020-Q1-E1-E8

Method BLASTX
NCBI GI g2842490
BLAST score 244
E value 7.0e-27
Match length 76
% identity 82

NCBI Description (AL021749) heat-shock protein [Arabidopsis thaliana]

Seq. No. 30

Contig ID 9 3.R1040

5'-most EST $j\overline{C}$ -gmst02400068a04d1

Method BLASTX
NCBI GI g2842490
BLAST score 151
E value 8.0e-10
Match length 38
% identity 71

NCBI Description (AL021749) heat-shock protein [Arabidopsis thaliana]

Seq. No. 31

Contig ID 10 1.R1040

5'-most EST LIB3139-004-P1-N1-H2

Method BLASTX
NCBI GI g3243234
BLAST score 1356
E value 1.0e-150



Match length % identity 84

(AF071477) isoflavone reductase related protein [Pyrus NCBI Description

32 Seq. No.

11 1.R1040 Contig ID

uC-gmrominsoy081c11b1 5'-most EST

BLASTN Method q166421 NCBI GI BLAST score 287 E value 1.0e-160 Match length 475

90 % identity

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

Seq. No. 33

11 2.R1040 Contig ID

LIB3139-094-P1-N1-A2 5'-most EST

Method BLASTN q166421 NCBI GI BLAST score 282 E value 1.0e-157 630 Match length % identity 89

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

34 Seq. No.

11 3.R1040 Contig ID 5'-most EST wvk700684812.h1

Method BLASTX NCBI GI q136636 BLAST score 222 4.0e-36 E value 84 Match length 94 % identity

UBIQUITIN-CONJUGATING ENZYME E2-17 KD 1 (UBIQUITIN-PROTEIN NCBI Description

LIGASE 1) (UBIQUITIN CARRIER PROTEIN 1)

>gi 1076424 pir S43781 ubiquitin-conjugating enzyme UBC1 -Arabidopsis thaliana >gi_442594_pdb_1AAK_ Ubiquitin Conjugating Enzyme (E.C.6.3.2.19) >gi_2981894_pdb_2AAK_ Ubiquitin Conjugating Enzyme From Arabidopsis Thaliana >gi 166924 (M62721) ubiquitin carrier protein [Arabidopsis thaliana] >gi 431260 (L19351) ubiquitin conjugating enzyme

[Arabidopsis thaliana]

Seq. No. 35

11 4.R1040 Contiq ID

5'-most EST LIB3106-009-Q1-K1-G1

Seq. No. 36

11 5.R1040 Contig ID

5'-most EST LIB3087-001-Q1-K1-E4

Method BLASTN NCBI GI g166421



BLAST score 8.0e-92 E value 248 Match length % identity 92

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

37 Seq. No.

11 6.R1040 Contig ID

LIB3092-033-Q1-K1-A7 5'-most EST

BLASTN Method g166421 NCBI GI 77 BLAST score E value 3.0e-35 Match length 125 90 % identity

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

38 Seq. No.

11 8.R1040 Contig ID

LIB3051-088-Q1-K1-B12 5'-most EST

BLASTN Method g166421 NCBI GI BLAST score 108 E value 2.0e-53 148 Match length 93 % identity

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

39 Seq. No.

11 9.R1040 Contig ID

5'-most EST LIB3106-105-Q1-K1-F5

BLASTN Method g166421 NCBI GI BLAST score 164 E value 4.0e-87 Match length 320 % identity 88

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

Seq. No. 40

12 1.R1040 Contig ID leu701146018.hl 5'-most EST

Method BLASTX g1717950 NCBI GI BLAST score 1036 1.0e-113 E value 242 Match length 81 % identity

UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT 2 NCBI Description PRECURSOR (RIESKE IRON-SULFUR PROTEIN 2) (RISP2) >gi 530052

(L16810) Rieske iron-sulfur protein [Nicotiana tabacum]

Seq. No. 41



12 2.R1040 Contig ID

LIB3106-095-Q1-K1-D10 5'-most EST

Method BLASTX NCBI GI g1717950 BLAST score 397 2.0e-38 E value 118 Match length % identity 68

UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT 2 NCBI Description

PRECURSOR (RIESKE IRON-SULFUR PROTEIN 2) (RISP2) >gi_530052

(L16810) Rieske iron-sulfur protein [Nicotiana tabacum]

42 Seq. No.

12 3.R1040 Contig ID 5'-most EST leu701151219.h1

43 Seq. No.

13 1.R1040 Contig ID

CPR2864_700208152_FL 5'-most EST

BLASTN Method g1673455 NCBI GI 443 BLAST score 0.0e+00E value 882 Match length 98 % identity

Z.mays rubisco small subunit gene NCBI Description

44 Seq. No.

14 1.R1040 Contig ID

CPR6566 700985454 FL 5'-most EST

Method BLASTX g2618686 NCBI GI BLAST score 1213 1.0e-134 E value

Match length 319 70 % identity

(AC002510) hypothetical protein [Arabidopsis thaliana] NCBI Description

45 Seq. No.

14 2.R1040 Contig ID

CPR6559_700731265_FL 5'-most EST

BLASTX Method NCBI GI g2618686 BLAST score 1017 E value 1.0e-121 310 Match length % identity

(AC002510) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 46

14 3.R1040 Contig ID

LIB3170-072-Q1-J1-E12 5'-most EST

BLASTX Method g2618686 NCBI GI BLAST score 309 4.0e-28 E value 88 Match length



% identity 65
NCBI Description (AC002510) hypothetical protein [Arabidopsis thaliana]

Seq. No. 47

Contig ID 15_1.R1040

5'-most EST fC-gmle7000742678f1

Method BLASTX
NCBI GI g2618686
BLAST score 989
E value 1.0e-107
Match length 389
% identity 48

NCBI Description (AC002510) hypothetical protein [Arabidopsis thaliana]

Seq. No. 48

Contig ID 15_2.R1040

5'-most EST CPR6560 700742678 FL

Method BLASTX
NCBI GI g3882018
BLAST score 416
E value 5.0e-66
Match length 324
% identity 47

NCBI Description (Y18377) cytokinin oxidase [Zea mays]

Seq. No. 49

Contig ID 16 1.R1040

5'-most EST uC-gmronoir010b05b1

Method BLASTN
NCBI GI g1663536
BLAST score 219
E value 1.0e-119
Match length 287
% identity 94

NCBI Description Glycine max clone RLG1 disease resistance protein homolog

gene, partial cds

Seq. No. 50

Contig ID 16_2.R1040 5'-most EST sat701011581.h1

Method BLASTN
NCBI GI g1663536
BLAST score 346
E value 0.0e+00
Match length 446
% identity 94

NCBI Description Glycine max clone RLG1 disease resistance protein homolog

gene, partial cds

Seq. No. 51

Contig ID 17 1.R1040

5'-most EST LIB3051-054-Q1-K2-C3

Method BLASTX
NCBI GI 94262226
BLAST score 277
E value 5.0e-24
Match length 195

% identity 46
NCBI Description (AC006200) putative protein kinase [Arabidopsis thaliana]

Seq. No. 52

Contig ID 17_2.R1040

5'-most EST LIB3139-085-P1-N1-A6

Method BLASTX
NCBI GI 94262226
BLAST score 197
E value 5.0e-15
Match length 128
% identity 45

NCBI Description (AC006200) putative protein kinase [Arabidopsis thaliana]

Seq. No. 53

Contig ID 17_3.R1040 5'-most EST lus701015877.h1

Seq. No. 54

Contig ID 19 1.R1040

5'-most EST LIB3030-011-Q1-B1-D10

Method BLASTX
NCBI GI g4325041
BLAST score 2609
E value 0.0e+00
Match length 668
% identity 80

NCBI Description (AF117339) FtsH-like protein Pftf precursor [Nicotiana

tabacum]

Seq. No. 55

Contig ID 19 2.R1040

5'-most EST jC-gmle01810022f02d1

Method BLASTN
NCBI GI g3808100
BLAST score 80
E value 6.0e-37
Match length 244
% identity 83

NCBI Description Capsicum annuum mRNA for chloroplast protease (CACP) from

the AAA atpase family

Seq. No. 56

Contig ID 19_3.R1040 5'-most EST ssr700559811.h1

Method BLASTX
NCBI GI g3808101
BLAST score 392
E value 7.0e-38
Match length 165
% identity 55

NCBI Description (AJ012165) chloroplast protease [Capsicum annuum]

Seq. No. 57

Contig ID 19_4.R1040

5'-most EST jC-gmf102220072h01a1



```
Seq. No.
                  20_1.R1040
Contig ID
                  CPR3447 700959152 FL
5'-most EST
Method
                  BLASTX
NCBI GI
                  `g1935000
BLAST score
                  404
E value
                  1.0e-38
Match length
                  238
% identity
                  35
NCBI Description (Y12314) GTPase [Schizosaccharomyces pombe]
                  59
Seq. No.
                  21 1.R1040
Contig ID
                  LIB3109-017-Q1-K1-B1
5'-most EST
Method
                  BLASTX
                  g3493172
NCBI GI
BLAST score
                  1329
                  1.0e-147
E value
                  313
Match length
% identity
                  81
NCBI Description (U89609) fiber annexin [Gossypium hirsutum]
                   60
Seq. No.
                  21_2.R1040
Contig ID
5'-most EST
                   fC-gmse700660534z1
Method
                  BLASTX
                   g3493172
NCBI GI
                   383
BLAST score
                   1.0e-36
E value
Match length
                   108
% identity
                   71
NCBI Description (U89609) fiber annexin [Gossypium hirsutum]
                   61
Seq. No.
                   21 4.R1040
Contig ID
5'-most EST
                   LIB3107-060-Q1-K1-G6
Method
                   BLASTX
                   q3493172
NCBI GI
                   424
BLAST score
E value
                   8.0e-42
                   108
Match length
% identity
NCBI Description (U89609) fiber annexin [Gossypium hirsutum]
                   62
Seq. No.
                   22 1.R1040
Contig ID
                   CPR3456_700739290_FL
5'-most EST
Method
                   BLASTX
                   g2213598
NCBI GI
BLAST score
                   436
                   2.0e-42
E value
```

332 Match length % identity 22

(AC000348) T7N9.18 [Arabidopsis thaliana] NCBI Description

63 Seq. No.

23 1.R1040 Contig ID

NCBI Description

thaliana]

```
CPR3463_700797953_FL
5'-most EST
                  BLASTX
Method
                   q3947733
NCBI GI
                   378
BLAST score
                   4.0e-36
E value
Match length
                   155
% identity
                   54
NCBI Description (AJ009719) NL25 [Solanum tuberosum]
Seq. No.
                   24 1.R1040
Contig ID
                   jex700908154.hl
5'-most EST
                   BLASTX
Method
                   g2088653
NCBI GI
                   1230
BLAST score
E value
                   1.0e-135
Match length
                   443
% identity
                   54
                   (AF002109) Hs1pro-1 related protein isolog [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   65
                   24 2.R1040
Contig ID
                   uC-gmrominsoy256f04b1
5'-most EST
Method
                   BLASTX
                   g2088653
NCBI GI
                   390
BLAST score
E value
                   2.0e-37
Match length
                   130
                   59
% identity
                   (AF002109) Hs1pro-1 related protein isolog [Arabidopsís
NCBI Description
                   thaliana]
                   66
 Seq. No.
                   24_3.R1040
 Contig ID
                   fC-gmle700557849f3
 5'-most EST
                   BLASTX
 Method
                   g2088653
 NCBI GI
                   413
 BLAST score
                   4.0e-40
 E value
                   123
 Match length
                   65
 % identity
                    (AF002109) Hs1pro-1 related protein isolog [Arabidopsis
 NCBI Description
                   thaliana]
                    67
 Seq. No.
                    24 5.R1040
 Contig ID
                    fua701040369.hl
 5'-most EST
 Method
                    BLASTX
                    g2088653
 NCBI GI
 BLAST score
                    352
                    3.0e-33
 E value
                    134
 Match length
                    52
 % identity
```

(AF002109) Hs1pro-1 related protein isolog [Arabidopsis

```
Seq. No.
                   25 1.R1040
Contig ID
5'-most EST
                   CPR3468 700835511 FL
Method
                   BLASTX
NCBI GI
                   g3947735
                   499
BLAST score
                   1.0e-49
E value
Match length
                   480
% identity
                   41
                  (AJ009720) NL27 [Solanum tuberosum]
NCBI Description
Seq. No.
                   69
                   26 1.R1040
Contig ID
                   fC-gmro700562665z1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3947735
BLAST score
                   670
                   5.0e-70
E value
Match length
                   358
                   41
% identity
NCBI Description
                  (AJ009720) NL27 [Solanum tuberosum]
Seq. No.
                   70
                   26 2.R1040
Contig ID
5'-most EST
                   CPR3475 700870127 FL
Method
                   BLASTX
NCBI GI
                   q3860167
BLAST score
                   322
E value
                   6.0e-29
Match length
                   358
                   26
% identity
NCBI Description
                  (AF098964) disease resistance protein RPP1-WsC [Arabidopsis
                   thaliana]
                   71
Seq. No.
Contig ID
                   26 3.R1040
5'-most EST
                   CPR10175 FL
Method
                   BLASTX
NCBI GI
                   q3947735
BLAST score
                   442
E value
                   3.0e-43
Match length
                   360
% identity
NCBI Description (AJ009720) NL27 [Solanum tuberosum]
Seq. No.
                   72
Contig ID
                   26 4.R1040
5'-most EST
                   pcp700990320.hl
```

Seq. No. 73

Contig ID 26_5.R1040

5'-most EST LIB3028-048-Q1-B1-A10

Method BLASTX
NCBI GI 94204265
BLAST score 158
E value 2.0e-10
Match length 140

```
% identity
                  (AC005223) 45643 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   26 6.R1040
Contig ID
                   LIB3053-011-Q1-N1-A10
5'-most EST
Method
                   BLASTX
                   q4388834
NCBI GI
BLAST score
                   333
                   7.0e-31
E value
Match length
                   199
                   18
% identity
                   (ACO06528) putative disease resistance protein RPP1, 3'
NCBI Description
                   partial [Arabidopsis thaliana]
                   75
Seq. No.
                   26 11.R1040
Contig ID
5'-most EST
                   dpv701099972.h1
Seq. No.
                   26 14.R1040
Contig ID
5'-most EST
                   zs<del>q</del>701118496.h1
Seq. No.
                   27 1.R1040
Contig ID
                   CPR3479_700891859_FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1086263
BLAST score
                   384
E value
                   2.0e-36
Match length
                   412
                   32
% identity
                   TMV resistance protein N - tobacco (Nicotiana glutinosa)
NCBI Description
                   >gi 558887 (U15605) N [Nicotiana glutinosa]
                   78
Seq. No.
                   28 1.R1040
Contig ID
                   leu701154138.h1
5'-most EST
Method
                   BLASTX
                   g2213598
NCBI GI
                   403
BLAST score
                   1.0e-38
E value
Match length
                   342
                   21
% identity
                   (AC000348) T7N9.18 [Arabidopsis thaliana]
NCBI Description
                   79
Seq. No.
```

Contig ID 30_1.R1040
5'-most EST gsv701053417.h1
Method BLASTX
NCBI GI g2316016
BLAST score 704
E value 3.0e-74
Match length 225
% identity 35

% identity 35
NCBI Description (U92650) MRP-like ABC transporter [Arabidopsis thaliana]

```
Seq. No.
Contig ID
                  30 2.R1040
5'-most EST
                  LIB3138-032-Q1-N1-B2
Method
                  BLASTX
                  g2316016
NCBI GI
BLAST score
                  353
E value
                  2.0e-33
Match length
                  114
% identity
                  61
NCBI Description (U92650) MRP-like ABC transporter [Arabidopsis thaliana]
Seq. No.
                  81
Contig ID
                  34 1.R1040
5'-most EST
                  CPR6213 700991711 FL
                  BLASTX
Method
NCBI GI
                  q2088653
BLAST score
                  1154
E value
                  1.0e-126
Match length
                  434
                  53
% identity
NCBI Description
                  (AF002109) Hslpro-1 related protein isolog [Arabidopsis
                  thaliana]
```

Seq. No. 82

Contig ID 34_2.R1040

5'-most EST LIB3049-052-Q1-E1-C5

Method BLASTX
NCBI GI g2088653
BLAST score 586
E value 1.0e-121
Match length 435
% identity 53

NCBI Description (AF002109) Hs1pro-1 related protein isolog [Arabidopsis

thaliana]

Seq. No. 83

Contig ID 37_1.R1040

Method BLASTX
NCBI GI g4490297
BLAST score 605
E value 1.0e-86
Match length 617
% identity 36

NCBI Description (AL035678) putative protein [Arabidopsis thaliana]

Seq. No. 84

Contig ID 37 2.R1040

5'-most EST uC-gmronoir032h02b1

Seq. No. 85

Contig ID 37_3.R1040

5'-most EST uC-gmropic011e09b1

Seq. No. 86

Contig ID 39_1.R1040



Method BLASTX
NCBI GI g3600040
BLAST score 479
E value 3.0e-47
Match length 402
% identity 36

NCBI Description (AF080119) similar to Arabidopsis thaliana disease

resistance protein RPS2 (GB:U14158) [Arabidopsis thaliana]

Seq. No. 87

Contig ID 41_1.R1040

Method BLASTX
NCBI GI g2213598
BLAST score 408
E value 1.0e-39
Match length 158
% identity 28

NCBI Description (AC000348) T7N9.18 [Arabidopsis thaliana]

Seq. No. 88

Contig ID 42 1.R1040

5'-most EST LIB3051-038-Q1-K1-E12

Method BLASTX
NCBI GI g4580460
BLAST score 2010
E value 0.0e+00
Match length 446
% identity 90

NCBI Description (AC006081) putative 26S Protease Subunit 4 [Arabidopsis

thaliana]

Seq. No. 89

Contig ID 42_2.R1040 5'-most EST gsv701051332.h1

Method BLASTX
NCBI GI g4580460
BLAST score 622
E value 6.0e-65
Match length 127
% identity 94

NCBI Description (AC006081) putative 26S Protease Subunit 4 [Arabidopsis

thaliana]

Seq. No. 90

Contig ID 42 3.R1040

5'-most EST LIB3051-047-Q1-K1-D7

Method BLASTN
NCBI GI g556557
BLAST score 122
E value 6.0e-62
Match length 302
% identity 85

NCBI Description Rice mRNA for homologue of Tat binding protein, complete

cds

Seq. No. 91

```
Contig ID
                   45 1.R1040
5'-most EST
                   zsq701117822.h2
Method
                   BLASTX
NCBI GI
                   q3947735
BLAST score
                   424
E value
                   5.0e-41
Match length
                   435
% identity
                   40
NCBI Description
                  (AJ009720) NL27 [Solanum tuberosum]
```

Seq. No. 92

Contig ID 45_2.R1040

5'-most EST LIB3139-050-P1-N1-E6

Method BLASTX
NCBI GI g730125
BLAST score 1929
E value 0.0e+00
Match length 488
% identity 73

NCBI Description NADPH-CYTOCHROME P450 REDUCTASE >gi_322739_pir_S31502

NADPH--ferrihemoprotein reductase (EC 1.6.2.4) - Madagascar

periwinkle >gi_18139_emb_CAA49446_ (X69791)

NADPH--ferrihemoprotein reductase [Catharanthus roseus]

Seq. No. 93

Contig ID 45_3.R1040

5'-most EST jC-gmro02910024d12a1

Method BLASTX
NCBI GI g2580499
BLAST score 609
E value 1.0e-112
Match length 323
% identity 66

NCBI Description (U67186) NADPH: ferrihemoprotein oxidoreductase

[Eschscholzia californica]

Seq. No. 94

Contig ID 45 4.R1040 5'-most EST g5058215 Method BLASTX NCBI GI g2580499 BLAST score 220 1.0e-17 E value Match length 98 % identity 45

NCBI Description (U67186) NADPH: ferrihemoprotein oxidoreductase

[Eschscholzia californica]

Seq. No. 95

Contig ID 45 5.R1040

5'-most EST LIB3051-067-Q1-K1-H3

Method BLASTX
NCBI GI g480669
BLAST score 366
E value 5.0e-35
Match length 106
% identity 68





NCBI Description NADPH--ferrihemoprotein reductase (EC 1.6.2.4) - Jerusalem artichoke (fragment) >gi 1359894 emb CAA81210 (Z26251) NADPH-ferrihemoprotein reductase [Helianthus tuberosus] Seq. No. Contig ID 45 6.R1040 5'-most EST LIB3170-037-Q1-K1-D8 Method BLASTX NCBI GI q730125 BLAST score 312 E value 3.0e-28 Match length 176 % identity 73 NADPH-CYTOCHROME P450 REDUCTASE >gi_322739_pir__S31502 NCBI Description NADPH--ferrihemoprotein reductase (EC 1.6.2.4) - Madagascar periwinkle >gi_18139_emb_CAA49446_ (X69791) NADPH--ferrihemoprotein reductase [Catharanthus roseus] 97 Seq. No. 45 7.R1040 Contig ID $vz\overline{y}700755269.h1$ 5'-most EST Method BLASTX NCBI GI q99737 BLAST score 542 E value 1.0e-55 Match length 150 % identity NCBI Description NADPH--ferrihemoprotein reductase (EC 1.6.2.4) -Arabidopsis thaliana >gi 16189 emb CAA46815 (X66017) NADPH-ferrihemoprotein reductase [Arabidopsis thaliana] 98 Seq. No. Contig ID 46 1.R1040 5'-most EST uC-gmropic067a07b1 Method BLASTX NCBI GI q322752 BLAST score 1146 E value 0.0e+00427 Match length 76 % identity auxin-independent growth promoter - Nicotiana tabacum NCBI Description >gi 559921 emb CAA56570 (X80301) axi 1 [Nicotiana tabacum] Seq. No. 99 48 1.R1040 Contig ID 5'-most EST CPR6563 700945759 FL Method BLASTX

NCBI GI g2618686 BLAST score 671 3.0e-70E value Match length 220 % identity 60

NCBI Description (AC002510) hypothetical protein [Arabidopsis thaliana]

Seq. No. 100

Contig ID 49 1.R1040

5'-most EST CPR6564_700953259_FL



```
Method
                  BLASTX
NCBI GI
                  q4191780
BLAST score
                   281
                  9.0e-25
E value
                  95
Match length
                  54
% identity
                  (AC005917) putative cytokinin oxidase [Arabidopsis
NCBI Description
                  thaliana]
                  101
Seq. No.
                   51 1.R1040
Contig ID
5'-most EST
                  CPR6639_700684235_FL
Method
                  BLASTX
                   q2258317
NCBI GI
                   172
BLAST score
                   6.0e-12
E value
                   132
Match length
                   37
% identity
                  (AF004879) resistance complex protein I2C-2 [Lycopersicon
NCBI Description
                   esculentum]
                   102
Seq. No.
                   52_1.R1040
Contig ID
5'-most EST
                   CPR6640_700942632_FL
Method
                   BLASTX
                   g3947733
NCBI GI
                   218
BLAST score
E value
                   4.0e-17
Match length
                   98
                   47
% identity
NCBI Description (AJ009719) NL25 [Solanum tuberosum]
Seq. No.
                   103
                   54 1.R1040
Contig ID
5'-most EST
                   zsq701118302.h1
Method
                   BLASTN
                   q310562
NCBI GI
BLAST score
                   89
E value
                   7.0e-42
Match length
                   336
% identity
                   88
NCBI Description Glycine max calmodulin (SCaM-5) mRNA, complete cds
                   104
Seq. No.
                   54 2.R1040
Contig ID
                   g310562 FL
5'-most EST
Method
                   BLASTN
NCBI GI
                   g310562
```

BLAST score 873 0.0e + 00E value

873 Match length 100 % identity

NCBI Description Glycine max calmodulin (SCaM-5) mRNA, complete cds

105 Seq. No.

54 3.R1040 Contig ID 5'-most EST zzp700831855.h1



Method BLASTN
NCBI GI g310562
BLAST score 115
E value 4.0e-58
Match length 123
% identity 98

NCBI Description Glycine max calmodulin (SCaM-5) mRNA, complete cds

Seq. No. 106

Contig ID 55_1.R1040

5'-most EST fC-gmle700560625a1

Method BLASTX
NCBI GI g2853078
BLAST score 393
E value 7.0e-77
Match length 594
% identity 34

NCBI Description (AL021768) TMV resistance protein N-like [Arabidopsis

thaliana]

Seq. No. 107

Contig ID 55 2.R1040

5'-most EST fC-gmle700560625i1

Method BLASTX
NCBI GI g2853078
BLAST score 317
E value 6.0e-29
Match length 201
% identity 33

NCBI Description (AL021768) TMV resistance protein N-like [Arabidopsis

thaliana]

Seq. No. 108

Contig ID 56 1.R1040

5'-most EST uC-gmrominsoy185f08b1

Method BLASTX
NCBI GI g3947735
BLAST score 638
E value 4.0e-66
Match length 402
% identity 33

NCBI Description (AJ009720) NL27 [Solanum tuberosum]

Seq. No. 109

Contig ID 57_1.R1040

5'-most EST LIB3051-041-Q1-K1-C3

Method BLASTX
NCBI GI g3947733
BLAST score 162
E value 1.0e-10
Match length 112
% identity 34

NCBI Description (AJ009719) NL25 [Solanum tuberosum]

Seq. No. 110

Contig ID 58_1.R1040

```
BLASTX
Method
NCBI GI
                   a4092774
                   669
BLAST score
                   1.0e-69
E value
Match length
                   604
                   31
% identity
NCBI Description
                   (AF105140) disease resistance gene homolog 9N [Brassica
                   111
Seq. No.
Contig ID
                   59 1.R1040
5'-most EST
                   CPR6649_700999286_FL
Method
                   BLASTX
                   q3947735
NCBI GI
                   373
BLAST score
                   3.0e - 35
E value
```

301 Match length % identity 42

(AJ009720) NL27 [Solanum tuberosum] NCBI Description

112 Seq. No. Contig ID 60 1.R1040 5'-most EST zhf700962285.h1 Method BLASTX

NCBI GI q1652057 BLAST score 1192 1.0e-131 E value Match length 535 % identity 51

(D90902) hypothetical protein [Synechocystis sp.] NCBI Description

113 Seq. No.

Contig ID 60 2.R1040 sat701012774.hl 5'-most EST

114 Seq. No. 65 1.R1040 Contig ID q5606563 5'-most EST Method BLASTX g3080405 NCBI GI 353 BLAST score 4.0e-33

E value 83 Match length 77 % identity

(AL022603) Lsd1 like protein [Arabidopsis thaliana] NCBI Description

>gi 4455269 emb CAB36805.1 (AL035527) Lsd1 like protein

[Arabidopsis thaliana]

115 Seq. No.

66 1.R1040 Contig ID

LIB3106-074-Q1-K1-G6 5'-most EST

BLASTX Method g2738248 NCBI GI BLAST score 3400 E value 0.0e + 00Match length 762 % identity 85

83



```
(U97200) cobalamin-independent methionine synthase
NCBI Description
                  [Arabidopsis thaliana]
Seq. No.
                  116
                  66 2.R1040
Contig ID
                  epx701108427.h1
5'-most EST
                  BLASTX
Method
                  q1914683
NCBI GI
                  294
BLAST score
                  3.0e-26
E value
                  74
Match length
                  77
% identity
                  (Y12013) RAD23, isoform I [Daucus carota]
NCBI Description
Seq. No.
                  117
                  66 3.R1040
Contig ID
                  LIB3092-020-Q1-K1-C11
5'-most EST
                  BLASTX
Method
                  q1914683
NCBI GI
                  325
BLAST score
                  4.0e-30
E value
                  109
Match length
% identity
                  61
                  (Y12013) RAD23, isoform I [Daucus carota]
NCBI Description
                  118
Seq. No.
                   66 6.R1040
Contig ID
                  crh700854007.hl
5'-most EST
                   BLASTX
Method
                   g1914683
NCBI GI
                   194
BLAST score
                   6.0e-15
E value
                   52
Match length
                   75
% identity
                  (Y12013) RAD23, isoform I [Daucus carota]
NCBI Description
                   119
Seq. No.
                   66_7.R1040
Contig ID
                   pxt700943841.h1
5'-most EST
                   BLASTX
Method
                   a974782
NCBI GI
                   418
BLAST score
                   3.0e-46
E value
                   111
Match length
% identity
                   86
                   (Z49150) cobalamine-independent methionine synthase
NCBI Description
                   [Solenostemon scutellarioides]
```

 Seq. No.
 120

 Contig ID
 66_9.R1040

 5'-most EST
 LIB3050-019-Q1-K1-A1

 Method
 BLASTN

 NCBI GI
 g974781

 BLAST score
 125

 E value
 8.0e-64

Match length 313 % identity 87



NCBI Description C.blumei kinetoplast met gene for cobalamine-independent methionine synthase

Seq. No. 121

Contig ID 66_16.R1040 5'-most EST wvk700682724.h1

Method BLASTX
NCBI GI g2738248
BLAST score 201
E value 1.0e-31
Match length 82
% identity 90

NCBI Description (U97200) cobalamin-independent methionine synthase

[Arabidopsis thaliana]

Seq. No. 122

Contig ID 67_1.R1040 5'-most EST zsg701118017.h2

Method BLASTX
NCBI GI g4376158
BLAST score 1940
E value 0.0e+00
Match length 472
% identity 82

NCBI Description (X98873) aspartate kinase [Arabidopsis thaliana]

Seq. No. 123

Contig ID 67_2.R1040

5'-most EST CPR9211_700959057_FL

Method BLASTX

NCBI GI g4376158

BLAST score 1866

E value 0.0e+00

Match length 458

% identity 81

NCBI Description (X98873) aspartate kinase [Arabidopsis thaliana]

Seq. No. 124

Contig ID 67_3.R1040

5'-most EST uC-gmropic064c02b1

Method BLASTN
NCBI GI g2257742
BLAST score 115
E value 1.0e-57
Match length 375
% identity 83

NCBI Description Arabidopsis thaliana lysine-sensitive aspartate kinase

mRNA, complete cds

Seq. No. 125

Contig ID 69 1.R1040

Method BLASTN
NCBI GI g2970446
BLAST score 905
E value 0.0e+00
Match length 1341



% identity 99

NCBI Description Glycine max aspartokinase-homoserine dehydrogenase

(AK-HSDH) mRNA, complete cds

Seq. No. 126

Contig ID 69_2.R1040 5'-most EST dpv701100440.h1

Method BLASTX
NCBI GI g2970447
BLAST score 2096
E value 0.0e+00
Match length 417
% identity 99

NCBI Description (AF049706) aspartokinase-homoserine dehydrogenase [Glycine

max]

Seq. No. 127

Contig ID 69 3.R1040

5'-most EST fC-gmro700865466a2

Method BLASTN
NCBI GI g2104674
BLAST score 169
E value 1.0e-89
Match length 477
% identity 84

NCBI Description V.faba mRNA for transcription factor containing bZIP

Seq. No. 128

Contig ID 69 4.R1040

5'-most EST fC-gmro700798021b2

Method BLASTN
NCBI GI g2970446
BLAST score 648
E value 0.0e+00
Match length 656
% identity 100

NCBI Description Glycine max aspartokinase-homoserine dehydrogenase

(AK-HSDH) mRNA, complete cds

Seq. No. 129

Contig ID 69 5.R1040

5'-most EST fC-gmle700557507r2

Method BLASTN
NCBI GI g2970446
BLAST score 519
E value 0.0e+00
Match length 535
% identity 99

NCBI Description Glycine max aspartokinase-homoserine dehydrogenase

(AK-HSDH) mRNA, complete cds

Seq. No. 130

Contig ID 71_1.R1040 5'-most EST uxk700672773.h1

Method BLASTX NCBI GI g1762933 BLAST score 485

```
E value 2.0e-48
Match length 186
% identity 53
```

NCBI Description (U66263) tumor-related protein [Nicotiana tabacum]

Seq. No. 131
Contig ID 72 1.R1040

Method BLASTX
NCBI GI g2754816
BLAST score 374
E value 2.0e-35
Match length 176
% identity 43

NCBI Description (AF021346) non-race specific disease resistance protein

[Arabidopsis thaliana]

Seq. No. 132

Contig ID 74 1.R1040

5'-most EST LIB3051-114-Q1-K1-H12

Method BLASTX
NCBI GI g1742961
BLAST score 1892
E value 0.0e+00
Match length 505
% identity 73

NCBI Description (X94756) cystathionine gamma-synthase [Arabidopsis

thaliana]

Seq. No. 133

Contig ID 74 2.R1040

5'-most EST fC-gmse7000751425d1

Method BLASTX
NCBI GI g1791309
BLAST score 380
E value 2.0e-36
Match length 83
% identity 87

NCBI Description (U83500) cystathionine gamma-synthase [Arabidopsis

thaliana] >gi_2852454_dbj_BAA24699_ (AB010888)

cystathionine gamma-synthase [Arabidopsis thaliana]

Seq. No. 134

Contig ID 74_3.R1040

5'-most EST LIB3051-006-Q1-E1-G10

Method BLASTX
NCBI GI g1791309
BLAST score 424
E value 9.0e-42
Match length 146
% identity 61

NCBI Description (U83500) cystathionine gamma-synthase [Arabidopsis

thaliana] >gi_2852454_dbj_BAA24699_ (AB010888)

cystathionine gamma-synthase [Arabidopsis thaliana]

Seq. No. 135

Contig ID 74_10.R1040



eep700867801.h1 5'-most EST BLASTX Method

g3202028 NCBI GI 395 BLAST score 3.0e-38 E value

96 Match length 83 % identity

(AF069317) cystathionine gamma-synthase [Mesembryanthemum NCBI Description

crystallinum]

136 Seq. No.

75 1.R1040 Contig ID

CPR9221 700567374 FL 5'-most EST BLASTX Method NCBI GI g100728 BLAST score 1351 1.0e-149 E value

407 Match length 61 % identity

aspartate transaminase (EC 2.6.1.1) AAT3 precursor - proso NCBI Description

millet >gi 1084464 pir S53304 aspartate aminotransferase proso millet >gi_20601_emb_CAA45024_ (X63430) aspartate

aminotransferase [Panicum miliaceum]

>gi 435459_dbj_BAA04993_ (D25323) aspartate

aminotransferase [Panicum miliaceum]

137 Seq. No.

75 2.R1040 Contig ID

uC-gmropic106a04b1 5'-most EST

BLASTX Method g100728 NCBI GI 573 BLAST score 6.0e-59 E value Match length 128 % identity

NCBI Description aspartate transaminase (EC 2.6.1.1) AAT3 precursor - proso

millet >gi_1084464_pir__S53304 aspartate aminotransferase - proso millet >gi_20601_emb_CAA45024_ (X63430) aspartate

aminotransferase [Panicum miliaceum]

>gi 435459 dbj BAA04993 (D25323) aspartate

aminotransferase [Panicum miliaceum]

Seq. No.

Contig ID 78 1.R1040

5'-most EST CPR9214_LIB3039-051-A2_FL

Method BLASTX q2984139 NCBI GI BLAST score 789 E value 7.0e-84 Match length 338 % identity

(AE000760) aspartate-semialdehyde dehydrogenase [Aquifex NCBI Description

aeolicus]

Seq. No.

139

138

78 2.R1040 Contig ID 5'-most EST hrw701060760.h1



Method BLASTX
NCBI GI g416901
BLAST score 341
E value 7.0e-32
Match length 111
% identity 59
NCBI Description ASPARTA

ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (ASA DEHYDROGENASE) > gi_142828 (L08471) aspartate semialdehyde dehydrogenase [Bacillus subtilis] > gi_2634047_emb_CAB13548_ (Z99112) aspartate-semialdehyde dehydrogenase [Bacillus subtilis]

Seq. No. 140

Contig ID 79_1.R1040 5'-most EST ssr700556460.h1

Method BLASTX
NCBI GI g1705463
BLAST score 1460
E value 1.0e-162
Match length 373
% identity 75

NCBI Description BIOTIN SYNTHASE (BIOTIN SYNTHETASE) >gi_2129547_pir__S71201

biotin sythase - Arabidopsis thaliana $>gi_10453\overline{16}$ ($\overline{U2}4147$) biotin sythase [Arabidopsis thaliana] $>gi_1403662$ (U31806) BIO2 protein [Arabidopsis thaliana] $>gi_1769457$ (L34413)

biotin synthase [Arabidopsis thaliana] >gi_2288983

(AC002335) biotin synthase (Bio B) [Arabidopsis thaliana] >gi_1589016_prf__2209438A biotin synthase [Arabidopsis

thaliana]

Seq. No. 141

Contig ID 79 2.R1040

5'-most EST LIB3170-010-Q1-K1-G7

Seq. No. 142

Contig ID 79_3.R1040 5'-most EST wvk700685486.h1

Seq. No. 143

Contig ID 82_1.R1040 5'-most EST ssr700555656.h1

Method BLASTN
NCBI GI 94098966
BLAST score 335
E value 0.0e+00
Match length 602
% identity 96

NCBI Description Glycine max putative reistance gene analog genomic sequence

Seq. No. 144

Contig ID 82_2.R1040

5'-most EST LIB3093-028-Q1-K1-A3

Method BLASTX
NCBI GI g1086263
BLAST score 346
E value 4.0e-32
Match length 382
% identity 31



TMV resistance protein N - tobacco (Nicotiana glutinosa) >qi 558887 (U15605) N [Nicotiana glutinosa]

Seq. No. 145

NCBI Description

Contig ID 83 1.R1040

5'-most EST CPR9217_701001147_FL

Method BLASTX
NCBI GI g1708993
BLAST score 863
E value 7.0e-93
Match length 195
% identity 86

NCBI Description CYSTATHIONINE BETA-LYASE PRECURSOR (CBL) (BETA-CYSTATHIONASE) (CYSTEINE LYASE)

>gi_2129567_pir__S61429 cystathionine beta-lyase (EC
4.4.1.8) - Arabidopsis thaliana >gi_704397 (L40511)
cystathionine beta-lyase [Arabidopsis thaliana]

Seq. No. 146

Contig ID 84_1.R1040

5'-most EST uC-gmronoir063g05b1

Method BLASTX
NCBI GI g2129753
BLAST score 1559
E value 0.0e+00
Match length 466
% identity 85

NCBI Description threonine synthase (EC 4.2.99.2) precursor - Arabidopsis

thaliana (fragment) >gi_1448917 (L41666) threonine synthase

[Arabidopsis thaliana]

Seq. No. 147

Contig ID 85_1.R1040 5'-most EST ssr700556066.h1

Method BLASTX
NCBI GI g4490297
BLAST score 717
E value 3.0e-75
Match length 656
% identity 34

NCBI Description (AL035678) putative protein [Arabidopsis thaliana]

Seq. No. 148

Contig ID 85_2.R1040
5'-most EST g5688438

Method BLASTX

NCBI GI g4490297

BLAST score 270

E value 1.0e-23

Match length 129
% identity 41

NCBI Description (AL035678) putative protein [Arabidopsis thaliana]

Seq. No. 149

Contig ID 86_1.R1040 5'-most EST epx701108161.h1

Method BLASTX



NCBI GI q1709205 BLAST score 1051 E value 1.0e-115 268 Match length 74 % identity

NCBI Description MYO-INOSITOL-1(OR 4)-MONOPHOSPHATASE 3 (IMP 3) (INOSITOL MONOPHOSPHATASE 3) >gi_1098971 (U39059) myo-inositol

monophosphatase 3 [Lycopersicon esculentum]

Seq. No.

Contig ID 86 2.R1040

5'-most EST uC-gmflminsoy001g09b1

Method BLASTX NCBI GI g1709205 BLAST score 1092 E value 1.0e-119 Match length 268 % identity 77

MYO-INOSITOL-1(OR 4)-MONOPHOSPHATASE 3 (IMP 3) (INOSITOL NCBI Description

MONOPHOSPHATASE 3) >gi 1098971 (U39059) myo-inositol

monophosphatase 3 [Lycopersicon esculentum]

Seq. No. 151

Contig ID 86 3.R1040

5'-most EST LIB3092-041-Q1-K1-A9

Method BLASTX NCBI GI a3914097 208 BLAST score E value 5.0e-31 Match length 102 70 % identity

MYO-INOSITOL-1(OR 4)-MONOPHOSPHATASE (IMP) (INOSITOL NCBI Description

MONOPHOSPHATASE) >gi 2708322 (AF037220) inositol

monophosphatase [Mesembryanthemum crystallinum]

152 Seq. No.

Contig ID 87 1.R1040

5'-most EST CPR9728:2F701150047 FL

Method BLASTX NCBI GI g2245036 BLAST score 1287 1.0e-142 E value Match length 391 % identity 65

(Z97342) triacylglycerol lipase homolog [Arabidopsis NCBI Description

thaliana

Seq. No. 153

Contig ID 87 2.R1040 5'-most EST CPR9727 FL Method BLASTX NCBI GI a2245036 BLAST score 1289 E value 1.0e-142 Match length 381 % identity 65

NCBI Description (Z97342) triacylglycerol lipase homolog [Arabidopsis



thaliana]

```
Seq. No.
                   154
Contig ID
                   88 1.R1040
5'-most EST
                   CPR10281:2F700737217 FL
Method
                   BLASTN
NCBI GI
                   q476215
BLAST score
                   912
E value
                   0.0e + 00
Match length
                   1756
% identity
                  Glycine max Century 84 p24 oleosin isoform B gene, complete
NCBI Description
Seq. No.
                   155
Contig ID
                   88 2.R1040
5'-most EST
                   g5688194
Method
                   BLASTN
                   q944829
NCBI GI
BLAST score
                   856
                   0.0e+00
E value
Match length
                   860
                   100
% identity
NCBI Description G.max mRNA for 24 kDa oleosin isoform (clone P24/91)
Seq. No.
                   156
Contig ID
                   88 3.R1040
                   CPR10278:2F700678725 FL
5'-most EST
Method
                   BLASTN
NCBI GI
                   g476213
BLAST score
                   907
                   0.0e+00
E value
Match length
                   1025
                   97
% identity
                  Glycine max Century 84 p24 oleosin isoform A gene, complete
NCBI Description
                   cds
                   157
. Seq. No.
                   88 8.R1040
Contig ID
5'-most EST
                   zpv700759006.h1
Method
                   BLASTN
NCBI GI
                   g476213
BLAST score
                   172
E value
                   4.0e-92
Match length
                   276
% identity
                   91
                  Glycine max Century 84 p24 oleosin isoform A gene, complete
NCBI Description
```

Seq. No. 158

Contig ID 91_1.R1040

5'-most EST uC-gmflminsoy089b10b1

Method BLASTX
NCBI GI g4126401
BLAST score 1596
E value 1.0e-178
Match length 375



% identity 81
NCBI Description (AB011795) flavanone 3-hydroxylase [Citrus sinensis]

Seq. No. 159

Contig ID 92 1.R1040

5'-most EST fC-gmse700669702f4

Seq. No. 160

Contig ID 92_2.R1040 5'-most EST CPR10713_FL Method BLASTX NCBI GI g1332411 BLAST score 1410

E value 1.0e-157 Match length 331 % identity 78

NCBI Description (D85102) dihydroflavonol 4-reductase [Rosa hybrida]

Seq. No. 161

Contig ID 92_3.R1040

5'-most EST jC-gmst02400075d12d1

Method BLASTX
NCBI GI g1706377
BLAST score 250
E value 4.0e-21
Match length 79
% identity 66

NCBI Description DIHYDROFLAVONOL-4-REDUCTASE (DFR) (DIHYDROKAEMPFEROL

4-REDUCTASE) >gi_499018_emb_CAA53578_ (X75964)

dihydroflavonol reductase [Vitis vinifera]

Seq. No. 162

Contig ID 92_4.R1040 5'-most EST kl1701212272.h1

Method BLASTX
NCBI GI g1888485
BLAST score 657
E value 6.0e-69
Match length 164
% identity 76

NCBI Description (Y11749) dihydroflavonol 4-reductase [Vitis vinifera]

Seq. No. 163

Contig ID 92_5.R1040 5'-most EST jsh701065818.h1

Method BLASTX
NCBI GI g1888485
BLAST score 312
E value 7.0e-29
Match length 71
% identity 83

NCBI Description (Y11749) dihydroflavonol 4-reductase [Vitis vinifera]

Seq. No. 164

Contig ID 93_1.R1040 5'-most EST CPR10721_FL Method BLASTX



NCBI GI g1488255 BLAST score 1899 E value 0.0e+00 Match length 483 % identity 74

NCBI Description (U38416) ferulate-5-hydroxylase [Arabidopsis thaliana]

>gi_2961381_emb_CAA18128_ (AL022141) ferulate-5-hydroxylase

(FAH1) [Arabidopsis thaliana] >gi_3925365 (AF068574)

ferulate-5-hydroxylase [Arabidopsis thaliana]

Seq. No. 165

Contig ID 93_2.R1040

5'-most EST LIB3170-046-Q1-J1-B7 Method BLASTX

NCBI GI g1488255
BLAST score 553
E value 1.0e-56
Match length 125
% identity 82

NCBI Description (U38416) ferulate-5-hydroxylase [Arabidopsis thaliana]

>gi_2961381_emb_CAA18128_ (AL022141) ferulate-5-hydroxylase

(FAH1) [Arabidopsis thaliana] >gi_3925365 (AF068574)

ferulate-5-hydroxylase [Arabidopsis thaliana]

Seq. No. 166

Contig ID 93_3.R1040

5'-most EST uC-gmflminsoy082h04b1

Method BLASTX
NCBI GI g1488255
BLAST score 726
E value 4.0e-77
Match length 174
% identity 76

NCBI Description (U38416) ferulate-5-hydroxylase [Arabidopsis thaliana]

>gi_2961381_emb_CAA18128_ (AL022141) ferulate-5-hydroxylase

(FAH1) [Arabidopsis thaliana] >gi 3925365 (AF068574)

ferulate-5-hydroxylase [Arabidopsis thaliana]

Seq. No. 167

94 1.R1040 Contig ID 5'-most EST CPR10745 FL Method BLASTX NCBI GI g3660465 BLAST score 563 1.0e-57 E value Match length 191 % identity 60

NCBI Description (AJ001753) Inositol 1,3,4-Trisphosphate 5/6 kinase

[Arabidopsis thaliana]

Seq. No. 168

Contig ID 95_1.R1040

5'-most EST LIB3052-012-Q1-N1-D4

Method BLASTX
NCBI GI g3097321
BLAST score 1840
E value 0.0e+00



Match length 379 % identity 90

NCBI Description (AB013289) Bd 30K [Glycine max]

Seq. No. 169

Contig ID 95_6.R1040 5'-most EST txt700733563.h1

Method BLASTN
NCBI GI g3097320
BLAST score 264
E value 1.0e-147
Match length 357
% identity 97

NCBI Description Glycine max gene for Bd 30K, complete cds

Seq. No. 170

Contig ID 96_1.R1040
5'-most EST CPR10171_FL
Method BLASTN
NCBI GI g3452139
BLAST score 124
E value 2.0e-62
Match length 312

NCBI Description Glycine max mRNA for resistance protein, partial

Seq. No. 171

% identity

Contig ID 98_1.R1040 5'-most EST kll701215045.h1

88

Method BLASTX
NCBI GI g1213629
BLAST score 1711
E value 0.0e+00
Match length 480
% identity 68

NCBI Description (X95991) pectinesterase [Prunus persica]

Seq. No. 172

Contig ID 98_2.R1040 5'-most EST kl1701204315.h2

Method BLASTX
NCBI GI g1213629
BLAST score 220
E value 7.0e-18
Match length 91
% identity 52

NCBI Description (X95991) pectinesterase [Prunus persica]

Seq. No. 173

Contig ID 99 1.R1040

5'-most EST CPR10679:2F701202691H1 FL

Method BLASTX
NCBI GI g2832625
BLAST score 742
E value 2.0e-78
Match length 329
% identity 46



NCBI Description (AL021711) putative protein [Arabidopsis thaliana]

174 Seq. No. Contig ID 101 1.R1040 5'-most EST cpr9972 FL Method BLASTX q4100321 NCBI GI BLAST score 434 1.0e-42 E value 203 Match length

% identity 45 NCBI Description (U96642) A sunflower resistance gene homolog, partial cds

[Helianthus annuus]

Seq. No. 175

Contig ID 102 1.R1040 5'-most EST CPR2854 FL Method BLASTX q120663 NCBI GI BLAST score 1673 0.0e + 00E value Match length 429 78 % identity

NCBI Description GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE B PRECURSOR,

CHLOROPLAST >gi 66027 pir__DEPMNB

glyceraldehyde-3-phosphate dehydrogenase (NADP+)

(phosphorylating) (EC 1.2.1.13) B precursor, chloroplast - garden pea >gi_309671 (M55147) glyceraldehyde-3-phosphate

dehydrogenase B subunit [Pisum sativum]

Seq. No. 176

Contig ID 104 1.R1040

5'-most EST cpr6672:2F701214276_FL

Method BLASTX
NCBI GI g2258315
BLAST score 174
E value 6.0e-12
Match length 193
% identity 39

NCBI Description (AF004878) resistance complex protein I2C-1 [Lycopersicon

esculentum]

Seq. No. 177

Contig ID 105 1.R1040

Method BLASTX
NCBI GI g543616
BLAST score 344
E value 6.0e-32
Match length 213
% identity 40

NCBI Description kunitz type subtilisin inhibitor, CLSI-II - Canavalia

lineata

Seq. No. 178

Contig ID 106 4.R1040 5'-most EST leu701154480.h1



Method BLASTX
NCBI GI g541849
BLAST score 740
E value 2.0e-96
Match length 253
% identity 69
NCBI Description anthran
Arabido

anthranilate synthase (EC 4.1.3.27) beta chain - Arabidopsis thaliana >gi_403434 (L22585) anthranilate

synthase beta subunit [Arabidopsis thaliana]

Seq. No. 179

Contig ID 106 5.R1040

5'-most EST LIB3028-047-Q1-B1-C9

Method BLASTN
NCBI GI g1173641
BLAST score 837
E value 0.0e+00
Match length 1175
% identity 96

NCBI Description Glycine max lipoxygenase (lox7) mRNA, complete cds

Seq. No. 180

Contig ID 106 6.R1040 5'-most EST kl1701202328.h1

Method BLASTX
NCBI GI g1928991
BLAST score 306
E value 8.0e-28
Match length 90
% identity 73

NCBI Description (U92815) heat shock protein 70 precursor [Citrullus

lanatus]

Seq. No. 181

Contig ID 106_7.R1040 5'-most EST epx701108350.h1

Method BLASTX
NCBI GI g1236949
BLAST score 964
E value 0.0e+00
Match length 858
% identity 63

NCBI Description (U50075) lipoxygenase L-5 [Glycine max]

Seq. No. 182

106 8.R1040 Contig ID 5'-most EST g1173641 FL Method BLASTX NCBI GI g1173642 BLAST score 1190 0.0e + 00E value 854 Match length % identity 47

NCBI Description (U36191) lipoxygenase [Glycine max]

>gi_1588566_prf__2208476A lipoxygenase [Glycine max]

Seq. No. 183



106 9.R1040 Contig ID LIB3106-026-Q1-K1-H7 5'-most EST BLASTX Method g585418 NCBI GI 758 BLAST score 0.0e + 00E value 855 Match length 38 % identity

NCBI Description LIPOXYGENASE-4 (L-4) (VSP94) >gi_2160320_dbj_BAA03101_ (D13999) lipxygenase L-4 [Glycine max]

Seq. No. 184

Contig ID 108 1.R1040

5'-most EST jC-gmro02910022e06d1

Method BLASTX
NCBI GI g115608
BLAST score 1619
E value 0.0e+00
Match length 314
% identity 98

NCBI Description PHOSPHOENOLPYRUVATE CARBOXYLASE 1 (PEPCASE)

>gi_22563_emb_CAA33316_ (X15238) phosphoenolpyruvate

carboxylase (AA 1 - 970) [Zea mays]

Seq. No. 185

Contig ID 109 1.R1040

5'-most EST CPR10669:2F701207222H1_FL

Seq. No. 186

Contig ID 110_1.R1040
5'-most EST CPRI0386_FL
Method BLASTX
NCBI GI g2529678
BLAST score 402
E value 1.0e-38
Match length 190
% identity 31

NCBI Description (AC002535) unknown protein [Arabidopsis thaliana]

Seq. No. 187

Contig ID 112_1.R1040 5'-most EST kl1701212463.h1

Method BLASTX
NCBI GI g1619602
BLAST score 602
E value 4.0e-62
Match length 215
% identity 56

NCBI Description (Y08726) MtN3 [Medicago truncatula]

Seq. No. 188

Contig ID 112 3.R1040 5'-most EST kl1701205638.h1

Method BLASTX
NCBI GI g1619602
BLAST score 234
E value 1.0e-19



Match length 96 % identity 52

NCBI Description (Y08726) MtN3 [Medicago truncatula]

Seq. No. 189

Contig ID 113_1.R1040

5'-most EST 701060596H1.SEQ_FL

Method BLASTX
NCBI GI g134151
BLAST score 218
E value 2.0e-17
Match length 109
% identity 41

NCBI Description 3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1 (STEROID

5-ALPHA-REDUCTASE 1) (SR TYPE 1) >gi_1079480_pir__A55274 3-oxo-5alpha-steroid 4-dehydrogenase (EC 1.3.99.5) - human >gi 177767 (M32313) steroid 5-alpha-reductase (EC 1.3.99.5)

[Homo sapiens] >gi_338476 (M68886) steroid

5-alpha-reductase [Homo sapiens] >gi_3360435 (AF052126)

steroid 5-alpha-reductase [Homo sapiens]
>gi_4507201_ref_NP_001038.1_pSRD5A1_

steroid-5-alpha-reductase, alpha polypeptide 1 (3-oxo-5

alpha-steroid delta 4-dehydrogenase alpha 1)

Seq. No. 190

Contig ID 117 1.R1040 5'-most EST CPR9734 FL BLASTX Method NCBI GI q3135672 BLAST score 161 3.0e-10 E value 118 Match length % identity 39

NCBI Description (AF064070) putative 1-acyl-sn-glycerol-3-phosphate

acyltransferase [Burkholderia pseudomallei]

Seq. No. 191

Contig ID 117 2.R1040 5'-most EST leu701157888.h1

Seq. No. 192

Contig ID 118_1.R1040

5'-most EST 700732879H1.SEQ_FL

Method BLASTX
NCBI GI g3873678
BLAST score 162
E value 7.0e-11
Match length 79
% identity 42

NCBI Description (Z71178) Similarity with yeast hypothetical protein (Swiss

prot accession number P40526) [Caenorhabditis elegans]

Seq. No. 193

Contig ID 119 1.R1040 5'-most EST asn701132823.h1

Method BLASTX NCBI GI g3738284



```
BLAST score 448
E value 8.0e-44
Match length 239
% identity 46
```

NCBI Description (AC005309) unknown protein [Arabidopsis thaliana]

Seq. No. 194

Contig ID 119 2.R1040

5'-most EST jC-gmle01810049d01d1

Seq. No. 195

Contig ID 119 3.R1040

5'-most EST uC-gmrominsoy049e04b1

Seq. No. 196

Contig ID 119_4.R1040 5'-most EST zhf700964040.h1

Method BLASTX
NCBI GI g3738284
BLAST score 303
E value 2.0e-27
Match length 78
% identity 71

NCBI Description (AC005309) unknown protein [Arabidopsis thaliana]

Seq. No. 197

120 2.R1040 Contig ID $g55\overline{1}0071$ 5'-most EST Method BLASTX q1255954 NCBI GI BLAST score 239 E value 3.0e-26Match length 110 % identity 59

NCBI Description (Z70677) thioredoxin [Ricinus communis]

Seq. No. 198

120 3.R1040 Contig ID 5'-most EST q5687937 Method BLASTN NCBI GI g169972 BLAST score 1204 0.0e+00E value Match length 1744 % identity 96

NCBI Description Soybean glycinin A-la-B-x subunit mRNA, complete cds

Seq. No. 199

Contig ID 120 4.R1040

5'-most EST LIB3072-059-Q1-K1-B3

Method BLASTN
NCBI GI g218264
BLAST score 1281
E value 0.0e+00
Match length 1751
% identity 95

NCBI Description Glycine max mRNA for glycinin A2Bla subunit, complete cds



```
200
Seq. No.
                  122 1.R1040
Contig ID
                  LIB3050-003-Q1-E1-G6
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2511541
                  1505
BLAST score
                  1.0e-168
E value
                  392
Match length
                  76
% identity /
                  (AF020787) DNA-binding protein GBP16 [Oryza sativa]
NCBI Description
                   201
Seq. No.
                   122 2.R1040
Contig ID
                   jC-qmle01810085c10d1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4539457
BLAST score
                   613
                   3.0e-63
E value
                   308
Match length
                   49
% identity
                  (AL049500) heat shock transcription factor-like protein
NCBI Description
                   [Arabidopsis thaliana]
                   202
Seq. No.
                   122 3.R1040
Contig ID
                   fC-gmro700846171a1
5'-most EST
                   BLASTX
Method
                   g2511541
NCBI GI
BLAST score
                   909
                   2.0e-98
E value
                   251
Match length
% identity
                   77
NCBI Description (AF020787) DNA-binding protein GBP16 [Oryza sativa]
                   203
Seq. No.
                   122 4.R1040
Contig ID
                   awf700837002.hl
5'-most EST
                   BLASTX
Method
                   g4539457
NCBI GI
                   586
BLAST score
                   3.0e-60
E value
                   255
Match length
                   54
 % identity
NCBI Description (AL049500) heat shock transcription factor-like protein
                    [Arabidopsis thaliana]
                    204
 Seq. No.
                    122_6.R1040
 Contig ID
                   fC-gmro700864991g1
 5'-most EST
                   BLASTX
 Method
                    g2511541
 NCBI GI
 BLAST score
                    331
                    2.0e-30
 E value
 Match length
                    77
                    86
 % identity
```

NCBI Description (AF020787) DNA-binding protein GBP16 [Oryza sativa]



```
205
Seq. No.
                  122 7.R1040
Contig ID
5'-most EST
                  LIB3170-029-Q1-J1-G11
Method
                  BLASTX
NCBI GI
                  g3645985
BLAST score
                  482
                  3.0e-48
E value
Match length
                  84
% identity
                  (AL031581) 1-evidence=predicted by content;
NCBI Description
                  1-method=genefinder;084; 1-method score=23.36;
                  1-evidence end; 2-evidence=predicted by match;
                  2-match accession=AA141041;
                  2-match description=CK01110.3prime CK Drosophila
                  melanogaster
                  206
Seq. No.
                  122 8.R1040
Contig ID
5'-most EST
                  q4396625
Method
                  BLASTX
NCBI GI
                  g3645985
BLAST score
                  479
E value
                  4.0e-48
Match length
                  84
% identity
                  95
                  (AL031581) 1-evidence=predicted by content;
NCBI Description
                  1-method=genefinder;084; 1-method score=23.36;
                  1-evidence end; 2-evidence=predicted by match;
                  2-match_accession=AA141041;
                  2-match_description=CK01110.3prime CK Drosophila
                  melanogaster
Seq. No.
                  207
                  122 9.R1040
Contig ID
5'-most EST
                  2DC-01-Q1-B1-H1
                  BLASTX
Method
                  g3645985
NCBI GI
BLAST score
                  428
                  3.0e-42
E value
Match length
                  74
% identity
NCBI Description
                  (AL031581) 1-evidence=predicted by content;
                  1-method=genefinder;084; 1-method score=23.36;
                   1-evidence end; 2-evidence=predicted by match;
                  2-match accession=AA141041;
                  2-match description=CK01110.3prime CK Drosophila
                  melanogaster
                  208
Seq. No.
                  122 11.R1040
Contig ID
5'-most EST
                  LIB3028-028-Q1-B1-B8
                  BLASTX
Method
```

NCBI GI a3645985 BLAST score 7.0e-21E value Match length 41



```
% identity
                  (AL031581) 1-evidence=predicted by content;
NCBI Description
                  1-method=genefinder;084; 1-method_score=23.36;
                  1-evidence_end; 2-evidence=predicted by match;
                  2-match_accession=AA141041;
                  2-match description=CK01110.3prime CK Drosophila
                  melanogaster
Seq. No.
                  209
                  124 1.R1040
Contig ID
                  CPR10667:2F701214015H1_FL
5'-most EST
Method
                  BLASTX
                  q535454
NCBI GI
                  962
BLAST score
E value
                  1.0e-104
                  309
Match length
% identity
                  59
                  (U13940) cysteine proteinase [Alnus glutinosa]
NCBI Description
                  210
Seq. No.
                  124 2.R1040
Contig ID
                  LIB3050-026-Q1-K1-D9
5'-most EST
                  BLASTX
Method
                   g535454
NCBI GI
                                                 -000
                   238
BLAST score
                   5.0e-20
E value
                   94
Match length
                   51
% identity
NCBI Description (U13940) cysteine proteinase [Alnus glutinosa]
                   211
Seq. No.
                   124 3.R1040
Contig ID
                   kl1701211192.h1
5'-most EST
                   BLASTX
Method
                   g2944446
NCBI GI
BLAST score
                   177
                   3.0e-13
E value
                   51
Match length
                   71
% identity
NCBI Description (AF050756) cysteine endopeptidase precursor [Ricinus
                   communis]
                   212
Seq. No.
                   125 1.R1040
Contig ID
                   CPR10665:2F701213140H1_FL
5'-most EST
                   BLASTX
Method
                   g3688284
NCBI GI
                   1204
BLAST score
                   1.0e-132
E value
                   371
Match length
 % identity
                   (AJ011567) lanatoside 15'-O-acetylesterase [Digitalis
 NCBI Description
                   lanata]
```

Seq. No. 213

Contig ID 126_1.R1040

5'-most EST LIB3051-042-Q1-K1-H5



```
BLASTX
Method
                  g1946364
NCBI GI
                  1416
BLAST score
                  1.0e-157
E value
                  420
Match length
                  64
% identity
                  (U93215) lipase isolog [Arabidopsis thaliana]
NCBI Description
                  214
Seq. No.
                  127 1.R1040
Contig ID
                  CPR6969:2F700980490 FL
5'-most EST
                  BLASTX
Method
                   g4165132
NCBI GI
                   1783
BLAST score
                   0.0e + 00
E value
                   483
Match length
                   69
% identity
                  (AF098292) endo-beta-1,4-D-glucanase [Lycopersicon
NCBI Description
                   esculentum]
                   215
Seq. No.
                   127 2.R1040
Contig ID
                   LIB3109-002-Q1-K2-H3
5'-most EST
                   BLASTN
Method
                   g3170524
NCBI GI
                   104
BLAST score
                   3.0e-51
E value
                   224
Match length
                   87
% identity
NCBI Description Fragaria x ananassa cellulase (Cel2) mRNA, complete cds
                   216
Seq. No.
                   130 1.R1040
Contig ID
                   kl1701211539.h1
 5'-most EST
                   BLASTX
Method
                   g2367431
NCBI GI
                   1358
BLAST score
                   1.0e-150
E value
                   467
Match length
                   56
 % identity
NCBI Description (AF000403) putative cytochrome P450 [Lotus japonicus]
                   217
 Seq. No.
                   131 1.R1040
 Contig ID
                   CPR9001:2F700903466_FL
 5'-most EST
 Method
                   BLASTX
                   g3212880
 NCBI GI
 BLAST score
                   1185
 E value
                   1.0e-159
                   455
 Match length
 % identity
 NCBI Description (AC004005) putative Mlo protein [Arabidopsis thaliana]
```

218 Seq. No.

131_2.R1040 Contig ID pcp700989120.hl 5'-most EST

BLASTX Method



```
NCBI GI g2765817
BLAST score 404
E value 2.0e-39
Match length 135
% identity 67
NCBI Description (Z95352)
>gi_3892
```

(Z95352) AtMlo-h1 [Arabidopsis thaliana] >gi 3892049 gb AAC78258.1 AAC78258 (AC002330) AtMlo-h1

[Arabidopsis thaliana]

Seq. No. 219

Contig ID 131_4.R1040 5'-most EST fua701038423.h1

Method BLASTX
NCBI GI g2765817
BLAST score 291
E value 3.0e-26
Match length 88
% identity 62

NCBI Description (Z95352) AtMlo-h1 [Arabidopsis thaliana]

>gi_3892049_gb_AAC78258.1_AAC78258 (AC002330) AtMlo-h1

[Arabidopsis thaliana]

Seq. No. 220

Contig ID 132_1.R1040 5'-most EST leu701150403.h1

Method BLASTX
NCBI GI g2213598
BLAST score 389
E value 5.0e-37
Match length 312
% identity 16

NCBI Description (AC000348) T7N9.18 [Arabidopsis thaliana]

Seq. No. 221

Contig ID 133_1.R1040

5'-most EST fC-zmf1700343886_FL

Method BLASTX
NCBI GI g4580398
BLAST score 746
E value 6.0e-79
Match length 198
% identity 71

NCBI Description (AC007171) putative protein kinase APK1A [Arabidopsis

thaliana]

Seq. No. 222

Contig ID 134 1.R1040

5'-most EST fC-zmfl700344343_FL

Method BLASTX
NCBI GI g3482921
BLAST score 264
E value 2.0e-22
Match length 79
% identity 62

NCBI Description (AC003970) Unknown protein [Arabidopsis thaliana]

Seq. No. 223



```
135_1.R1040
Contig ID
                  fC-zmf1700352002_FL
5'-most EST
                  BLASTX
Method
NCBI GI
                  g585593
                  220
BLAST score
                  3.0e-17
E value
                  188
Match length
                  27
% identity
NCBI Description NUCLEAR PORE GLYCOPROTEIN P62 (NUCLEOPORIN P62)
                  >gi_432654_emb_CAA41411_ (X58521) nucleoporin p62 [Homo
                  sapiens]
                  224
Seq. No.
Contig ID
                  136 1.R1040
                  fC-zmle700422033_FL
5'-most EST
                  225
Seq. No.
                  137 1.R1040
Contig ID
                  fC-zmf1700351889 FL
5'-most EST
                  226
Seq. No.
                  138 1.R1040
Contig ID
5'-most EST
                  LIB3093-038-Q1-K1-F11
                  BLASTX
Method
NCBI GI
                  q4126473
                   876
BLAST score
                   3.0e-94
E value
                   211
Match length
                   81
% identity
                   (AB014884) adenylyl cyclase associated protein [Gossypium
NCBI Description
                  hirsutum]
                   227
Seq. No.
                   139 1.R1040
Contig ID
5'-most EST
                   fC-gmst700790741 FL
Method
                   BLASTN
NCBI GI
                   g1785614
                   49
BLAST score
E value
                   5.0e-18
Match length
                   126
% identity
                   85
                  Pisum sativum protein kinase homolog PsPK4 mRNA, complete
NCBI Description
                   cds
                   228
Seq. No.
                   139 2.R1040
Contig ID
                   uC-gmflminsoy058f07b1
5'-most EST
                   229
Seq. No.
                   139 3.R1040
```

Contig ID

LIB3094-030-Q1-K1-D3 5'-most EST

BLASTN Method g1480927 NCBI GI BLAST score 170 2.0e-90 E value 301 Match length % identity 89





Glycine max cdc2 protein kinase ((cdc2(1)) nonprocessed NCBI Description pseudogene Seq. No. 230 139 4.R1040 Contig ID LIB3094-072-Q1-K1-A6 5'-most EST BLASTN Method g1480927 NCBI GI 97 BLAST score 7.0e-47 E value 211 Match length 91 % identity Glycine max cdc2 protein kinase ((cdc2(1)) nonprocessed NCBI Description pseudogene Seq. No. 231 139 5.R1040 Contig ID jsh701064823.h1 5'-most EST BLASTN Method q1480927 NCBI GI 93 BLAST score 1.0e-44 E value Match length 137 92 % identity Glycine max cdc2 protein kinase ((cdc2(1)) nonprocessed NCBI Description pseudogene 232 Seq. No. 140 1.R1040 Contig ID fC-zmf1700344725 FL 5'-most EST BLASTX Method q3786005 NCBI GI 1535 BLAST score 1.0e-171 E value 365 Match length 80 % identity (AC005499) putative phosphoethanolamine NCBI Description cytidylyltransferase [Arabidopsis thaliana] 233 Seq. No. 141 1.R1040 Contig ID fC-zmf1700352293 FL 5'-most EST Method BLASTX g127112 NCBI GI BLAST score 450 4.0e-44E value 175 Match length % identity NCBI Description MAK16 PROTEIN >gi_73269_pir__BVBYK6 MAK16 protein - yeast

(Saccharomyces cerevisiae) >gi_171880 (J03852) MAK16 protein [Saccharomyces cerevisiae] >gi_595561 (U12980)

Mak16p: putative nuclear protein [Saccharomyces cerevisiae]

234 Seq. No. 142 1.R1040 Contig ID fC-gmse700645507_FL 5'-most EST BLASTN Method



```
g791097
NCBI GI
BLAST score
                  258
                  1.0e-143
E value
Match length
                  773
                  90
% identity
NCBI Description P.vulgaris plsB mRNA
```

235 Seq. No. 143 1.R1040 Contig ID

LIB3107-017-Q1-K1-H1 5'-most EST

BLASTX Method g1666096 NCBI GI BLAST score 1108 1.0e-121 E value Match length 308 % identity 69

(Y09113) dioxygenase [Marah macrocarpus] NCBI Description

236 Seq. No. Contig ID

144_1.R1040 fC-gmse700667947f1

5'-most EST Method BLASTX g3128225 NCBI GI BLAST score 1076 1.0e-117 E value

Match length 322 % identity 68

(AC004077) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No. 237

144 2.R1040 Contig ID

fC-gmse700674467a2 5'-most EST

BLASTX Method NCBI GI q3128225 170 BLAST score E value 8.0e-12 Match length 45 76 % identity

NCBI Description (AC004077) unknown protein [Arabidopsis thaliana]

Seq. No. 238

145 1.R1040 Contig ID

fC-zmse700838611 FL 5'-most EST

Method BLASTX g2252846 NCBI GI BLAST score 681 E value 7.0e-90 Match length 310 % identity 64

(AF013293) Similar to phospholipase D [Arabidopsis NCBI Description

thaliana]

239 Seq. No.

146 1.R1040 Contig ID

5'-most EST fC-zmse700838837 FL

Seq. No. 240



Contig ID 147_1.R1040 5'-most EST kl1701210137.h1 Method BLASTX NCBI GI q4510395

NCBI GI g4510395 BLAST score 1198 E value 1.0e-146 Match length 394 % identity 66

NCBI Description (AC006587) putative beta-galactosidase precursor

[Arabidopsis thaliana]

Seq. No. 241

Contig ID 148 1.R1040 5'-most EST zhf700960184.h1

Method BLASTX
NCBI GI g3914394
BLAST score 2455
E value 0.0e+00
Match length 559
% identity 83

NCBI Description 2,3-BISPHOSPHOGLYCERATE-INDEPENDENT PHOSPHOGLYCERATE MUTASE

(PHOSPHOGLYCEROMUTASE) (BPG-INDEPENDENT PGAM) (PGAM-I) >gi_2118335_pir__S60473 phosphoglycerate mutase (EC 5.4.2.1) - common ice plant >gi_602426 (U16021)

phosphoglyceromutase [Mesembryanthemum crystallinum]

Seq. No. 242

Contig ID 149_1.R1040 5'-most EST uaw700661032.h1

Method BLASTX
NCBI GI g2244744
BLAST score 347
E value 3.0e-32
Match length 122
% identity 57

NCBI Description (Y13676) bZIP DNA-binding protein [Antirrhinum majus]

Seq. No. 243

Contig ID 149 2.R1040

5'-most EST LIB3049-016-Q1-E1-B3

Method BLASTN
NCBI GI g394735
BLAST score 36
E value 1.0e-10
Match length 52
% identity 92

NCBI Description Rice lip19 mRNA for basic/leucine zipper protein

Seq. No. 244

Contig ID 149_3.R1040 5'-most EST zzp700836073.h1

Seq. No. 245

Contig ID 149 4.R1040

5'-most EST LIB3028-008-Q1-B1-A1

Seq. No. 246



150 1.R1040 Contig ID zsg701120848.h1 5'-most EST Method BLASTX NCBI GI g1362002 BLAST score 990 1.0e-107 E value 242 Match length 77 % identity protein kinase 1 - Arabidopsis thaliana >gi_166817 (L05561) NCBI Description protein kinase [Arabidopsis thaliana] 247 Seq. No. Contig ID 150 2.R1040 uC-gmropic058h06b1 5'-most EST Method BLASTX NCBI GI q1362002 BLAST score 256 9.0e-30 E value 130 Match length % identity 60 protein kinase 1 - Arabidopsis thaliana >gi_166817 (L05561) NCBI Description protein kinase [Arabidopsis thaliana] 248 Seq. No. 151 1.R1040 Contig ID kl1701214918.hl 5'-most EST BLASTX Method g2459435 NCBI GI BLAST score 1494 1.0e-166 E value Match length 452 % identity 63 (AC002332) putative serine carboxypeptidase [Arabidopsis NCBI Description thaliana] 249 Seq. No. 152 1.R1040 Contig ID 5'-most EST fC-gmro700562302_FL BLASTX Method g1888359 NCBI GI 294 BLAST score 6.0e-33 E value 179 Match length 69 % identity (Y11689) 3-ketoacyl-acyl carrier protein synthase III NCBI Description [Arabidopsis thaliana] 250 Seq. No. Contig ID 152 2.R1040 fC-gmse700854533a1 5'-most EST Method BLASTX q1706757 NCBI GI BLAST score 1048

1.0e-114 E value Match length 232 % identity

NCBI Description 3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE III A PRECURSOR





(BETA-KETOACYL-ACP SYNTHASE III A) (KAS III A) >gi_1276434 (U15935) beta-ketoacyl-ACP synthase III [Cuphea wrightii]

 Seq. No.
 251

 Contig ID
 153_1.R1040

 5'-most EST
 xpa700793632.h1

 Method
 BLASTX

 NCBI GI
 g2346978

 BLAST score
 344

 E value
 7.0e-32

NCBI Description (AB006601) ZPT2-14 [Petunia x hybrida]

Seq. No. 252

Match length

% identity

Contig ID 153_2.R1040

5'-most EST LIB3049-032-Q1-E1-B4

167

49

Method BLASTX
NCBI GI g2346976
BLAST score 356
E value 2.0e-33
Match length 152
% identity 54

NCBI Description (AB006600) ZPT2-13 [Petunia x hybrida]

Seq. No. 253

Contig ID 154_1.R1040 5'-most EST awf700836660.h1

Method BLASTX
NCBI GI g2664214
BLAST score 189
E value 2.0e-13
Match length 171
% identity 31

NCBI Description (AJ222646) G2484-1 [Arabidopsis thaliana]

Seq. No. 254

Contig ID 154 2.R1040

5'-most EST LIB3170-039-Q1-J1-H5

Seq. No. 255

Contig ID 155 1.R1040

5'-most EST fC-gmro700566983 FL

Seq. No. 256

Contig ID 156 1.R1040

Method BLASTX
NCBI GI g1346729
BLAST score 192
E value 4.0e-14
Match length 198
% identity 16

NCBI Description PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA >gi_886024 (U23820) protein kinase PkwA [Thermomonospora curvata]

Seq. No. 257



157 1.R1040 Contig ID 5'-most EST q5752858 Method BLASTX NCBI GI g2811029 BLAST score 385 6.0e-37 E value Match length 107 % identity 72

NCBI Description ACETYLORNITHINE AMINOTRANSFERASE PRECURSOR (ACOAT)

(ACETYLORNITHINE TRANSAMINASE) (AOTA)

>gi_1944511_emb_CAA69936_ (Y08680) acetylornithine

aminotransferase [Alnus glutinosa]

Seq. No. 258

Contig ID 158_1.R1040 5'-most EST fua701037507.h1

Method BLASTX
NCBI GI g1076318
BLAST score 383
E value 2.0e-36
Match length 109
% identity 41

NCBI Description dihydrolipoamide S-acetyltransferase (EC 2.3.1.12)

precursor - Arabidopsis thaliana (fragment)

>gi_559395_emb_CAA86300_ (Z46230) dihydrolipoamide acetyltransferase (E2) subunit of PDC [Arabidopsis

thaliana)

Seq. No. 259

Contig ID 158_2.R1040 5'-most EST zzp700833589.h1

Method BLASTX
NCBI GI g2129473
BLAST score 375
E value 1.0e-35
Match length 161
% identity 47

NCBI Description arabinogalactan-like protein - loblolly pine >gi_607774

(U09556) arabinogalactan-like protein [Pinus taeda]

Seq. No. 260

Contig ID 158_3.R1040 5'-most EST fua701041084.h1

Method BLASTX
NCBI GI g2129473
BLAST score 325
E value 5.0e-30
Match length 136
% identity 49

NCBI Description arabinogalactan-like protein - loblolly pine >gi_607774

(U09556) arabinogalactan-like protein [Pinus taeda]

Seq. No. 261

Contig ID 159 1.R1040

5'-most EST LIB3049-021-Q1-E1-A2

Method BLASTX NCBI GI q4115914



BLAST score 7.0e-53 E value Match length 324 % identity 38

(AF118222) contains similarity to Iron/Ascorbate family of NCBI Description oxidoreductases (Pfam: PF00671, Score=297.8, E=1.3e-85, N=1) [Arabidopsis thaliana] >gi 4539410 emb CAB40043.1 (AL049524) putative Fe(II)/ascorbate oxidase [Arabidopsis

thaliana]

Seq. No.

262 Contig ID 159 2.R1040 5'-most EST leu701153539.hl

Seq. No.

263

Contig ID

160 1.R1040

5'-most EST

LIB3028-040-Q1-B1-E7

Method BLASTX q2780955 NCBI GI BLAST score 238 E value 3.0e-19 Match length 214

% identity NCBI Description Phosphatidylinositol Transfer Protein Sec14p From

Saccharomyces Cerevisiae

Seq. No.

264

Contig ID

160 2.R1040

5'-most EST

LIB3049-026-Q1-E1-A3

Seq. No.

265

Contig ID

161 1.R1040

5'-most EST

jC~gmf102220098b09d1

Method BLASTX NCBI GI g1352186 BLAST score 1271 E value 1.0e-140 Match length 473 51

% identity

NCBI Description ALLENE OXIDE SYNTHASE PRECURSOR (HYDROPEROXIDE DEHYDRASE)

(CYTOCHROME P450 74) >gi 404866 (U00428) allene oxide

synthase [Linum usitatissimum]

Seq. No. 266

Contig ID 161 2.R1040

5'-most EST fC-gmse700751003d2

Method BLASTN g18662 NCBI GI BLAST score 903 E value 0.0e+00Match length 974 98 % identity

NCBI Description Glycine max hsp 70 gene

Seq. No.

267

161 3.R1040 Contig ID

5'-most EST

jC-gmst02400030e01d1

```
Method
                   g4263779
NCBI GI
                   540
BLAST score
                   4.0e-55
E value
                   147
Match length
                   67
% identity
                  (AC006068) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   268
Seq. No.
                   161 4.R1040
Contig ID
                   fC-qmro700793925r1
5'-most EST
                   BLASTX
Method
                   g4262242
NCBI GI
                   305
BLAST score
                   1.0e-27
E value
                   129
Match length
% identity
                   (AC006200) NADC homolog [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   161 5.R1040
Contig ID
                   zpv700760758.hl
5'-most EST
                   BLASTN
Method
                   q18662
NCBI GI
                   541
BLAST score
                   0.0e + 00
E value
                   729
Match length
                   94
% identity
NCBI Description Glycine max hsp 70 gene
                   270
Seq. No.
                   161 8.R1040
Contig ID
                   zsg701123403.h1
5'-most EST
                   BLASTN
Method
                   q18662
NCBI GI
                   150
BLAST score
                   6.0e-79
E value
                   202
Match length
                   94
 % identity
                  Glycine max hsp 70 gene
NCBI Description
                    271
 Seq. No.
                    162 1.R1040
 Contig ID
                   LIB3107-040-Q1-K1-F3
 5'-most EST
                    BLASTX
 Method
                    g2493494
 NCBI GI
                    1567
 BLAST score
                    1.0e-175
 E value
                    428
 Match length
                    65
```

>gi 619351_bbs_153537 CP-MII.2=serine carboxypeptidase [Hordeum vulgare=barley, cv. Alexis, aleurone, Peptide, 436

NCBI Description

% identity

272 Seq. No. Contig ID

aa]

163_1.R1040

SERINE CARBOXYPEPTIDASE II-2 PRECURSOR (CP-MII.2)



5'-most EST fC-zmst700621904_FL
Method BLASTX
NCBI GI g113029
BLAST score 1716

E value 0.0e+00 Match length 452 % identity 74

NCBI Description ISOCITRATE LYASE (ISOCITRASE) (ISOCITRATASE) (ICL)

>gi_68209_pir__WZCNIU isocitrate lyase (EC 4.1.3.1) upland cotton >gi_18486_emb_CAA36381_ (X52136) isocitrate

lyase (AA 1-576) [Gossypium hirsutum]

Seq. No. 273

Contig ID 164 1.R1040

5'-most EST fC-gmle700786490 FL

Method BLASTX
NCBI GI g4204849
BLAST score 948
E value 1.0e-102
Match length 223
% identity 78

NCBI Description (U55875) protein kinase [Arabidopsis thaliana]

Seq. No. 274

Contig ID 165 1.R1040

5'-most EST LIB3109-022-Q1-K1-G3

Method BLASTX
NCBI GI g3688123
BLAST score 2069
E value 0.0e+00
Match length 619
% identity 66

NCBI Description (AJ006293) granule-bound starch synthase [Antirrhinum

majus]

Seq. No. 275

Contig ID 165 5.R1040

5'-most EST LIB3092-018-Q1-K1-G2

Seq. No. 276

Contig ID 166 1.R1040

5'-most EST fC-zmro700835672_FL

Method BLASTX
NCBI GI g3450842
BLAST score 894
E value 3.0e-96
Match length 269
% identity 62

NCBI Description (AF080436) mitogen activated protein kinase kinase [Oryza

sativa]

Seq. No. 277

Contig ID 167 1.R1040

5'-most EST fC-zmro700835718_FL

Method BLASTX
NCBI GI g2829275
BLAST score 837



1.0e-89 E value Match length 236 % identity 70 NCBI Description

(AF044265) nucleoside diphosphate kinase 3 [Arabidopsis thaliana] >gi 3513740 (AF080118) contains similarity to nucleoside diphosphate kinases (Pfam: NDK.hmm, score: 301.12) [Arabidopsis thaliana] >gi_4539375_emb_CAB40069.1_

(AL049525) nucleoside diphosphate kinase 3 (ndpk3)

[Arabidopsis thaliana]

278 Seq. No.

Contig ID 167 2.R1040

5'-most EST LIB3093-031-Q1-K1-H3

Method BLASTX NCBI GI g2829275 BLAST score 125 E value 6.0e-09 Match length 70 % identity

NCBI Description (AF044265) nucleoside diphosphate kinase 3 [Arabidopsis

thaliana] >qi 3513740 (AF080118) contains similarity to nucleoside diphosphate kinases (Pfam: NDK.hmm, score: 301.12) [Arabidopsis thaliana] >gi 4539375_emb_CAB40069.1_

(AL049525) nucleoside diphosphate kinase 3 (ndpk3)

[Arabidopsis thaliana]

Seq. No. 279

Contig ID 168 1.R1040

5'-most EST fC-zmro700835944 FL

Method BLASTX NCBI GI g3901012 BLAST score 1162 E value 1.0e-128 Match length 274

% identity 75

(AJ130885) xyloglucan endotransglycosylase 1 [Fagus NCBI Description

sylvatica]

280 Seq. No.

Contig ID 168 2.R1040

5'-most EST uC-gmropic070h06b1

Method BLASTX NCBI GI g3901012 BLAST score 1197 1.0e-132 E value Match length 278 % identity

NCBI Description (AJ130885) xyloglucan endotransglycosylase 1 [Fagus

sylvatica]

Seq. No. 281

Contig ID 168 3.R1040

5'-most EST fC-gmse700670426g2

Method BLASTX NCBI GI q3901012 BLAST score 514 4.0e-52 E value



```
127
Match length
% identity
                   69
NCBI Description
                   (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
                   sylvatica]
                   282
Seq. No.
                   168_4.R1040
Contig ID
5'-most EST
                   zzp700832650.h1
Method
                   BLASTX
                   g3901012
NCBI GI
BLAST score
                   1119
E value
                   1.0e-123
                   249
Match length
                   80
% identity
                  (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
NCBI Description
                   sylvatica]
Seq. No.
                   283
Contig ID
                   168 5.R1040
5'-most EST
                   dpv701097331.h1
Method
                   BLASTX
                   g3901012
NCBI GI
BLAST score
                   327
                   2.0e-30
E value
Match length
                   71
% identity
                   87
                   (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
NCBI Description
                   sylvatica]
Seq. No.
                   284
                   169 1.R1040
Contig ID
5'-most EST
                   fua701039647.h1
Method
                   BLASTX
NCBI GI
                   g1915974
BLAST score
                   1420
E value
                   1.0e-158
                   319
Match length
                   86
% identity
NCBI Description
                   (U62329) fructokinase [Lycopersicon esculentum] >gi 2102693
                   (U64818) fructokinase [Lycopersicon esculentum]
                   285
Seq. No.
                   169 2.R1040
Contig ID
5'-most EST
                   uC-gmronoir063d06b1
Method
                   BLASTX
                   g585973
NCBI GI
BLAST score
                   756
E value
                   2.0e-80
Match length
                   182
                   77
% identity
                   FRUCTOKINASE >gi_626018_pir__S39997 fructokinase (EC 2.7.1.4) - potato >gi_297015_emb_CAA78283_ (Z12823)
NCBI Description
```

Contig ID 169 3.R1040

fructokinase [Solanum tuberosum]

fructokinase [Solanum tuberosum] >gi_1095321_prf__2108342A



```
LIB3055-004-Q1-N1-H3
5'-most EST
Method
                  BLASTX
                  q585973
NCBI GI
                  353
BLAST score
                  4.0e-53
E value
Match length
                  127
% identity
                  FRUCTOKINASE >gi 626018 pir__S39997 fructokinase (EC
NCBI Description
                  2.7.1.4) - potato >gi 297015_emb_CAA78283_ (Z12823)
                   fructokinase [Solanum tuberosum] >gi_1095321_prf__2108342A
                   fructokinase [Solanum tuberosum]
                   287
Seq. No.
                   170 1.R1040
Contig ID
                  fC-zmle700426440_FL
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2736286
BLAST score
                   322
                   1.0e-51
E value
                   125
Match length
% identity
                  (AF031079) isopentenyl diphosphate isomerase I [Camptotheca
NCBI Description
                   acuminata]
                   288
Seq. No.
                   171 5.R1040
Contig ID
5'-most EST
                   LIB3106-113-Q1-K1-C5
                   BLASTN
Method
                   q18730
NCBI GI
BLAST score
                   85
                   7.0e-40
E value
Match length
                   89
% identity
                   99
NCBI Description Soybean 4.5 - 5S rRNA intergenic region
Seq. No.
                   289
                   171 6.R1040
Contig ID
                   wvk700685569.h1
5'-most EST
                   BLASTN
Method
                   q18730
NCBI GI
BLAST score
                   72
                   3.0e-32
E value
Match length
                   72
                   100
% identity
                  Soybean 4.5 - 5S rRNA intergenic region
NCBI Description
                   290
Seq. No.
                   171 7.R1040
Contig ID
                   wrg700791869.hl
5'-most EST
                   291
 Seq. No.
                   171 8.R1040
 Contig ID
 5'-most EST
                   LIB3092-010-Q1-K1-F5
```

118

BLASTX

237 1.0e-19

q1173218

Method

NCBI GI BLAST score

E value



Match length 46
% identity 96
NCBI Description 40S RIBOSOMAL PROTEIN S15A >gi_440824 (L27461) ribosomal protein S15 [Arabidopsis thaliana] >gi_2150130 (AF001412) cytoplasmic ribosomal protein S15a [Arabidopsis thaliana]

Seq. No. 292

Contig ID 171 9.R1040

5'-most EST LIB3107-079-Q1-K1-G11

Method BLASTN
NCBI GI g18730
BLAST score 85
E value 7.0e-40
Match length 97
% identity 97

NCBI Description Soybean 4.5 - 5S rRNA intergenic region

Seq. No. 293

Contig ID 171_10.R1040

5'-most EST LIB3049-047-Q1-E1-C11

Seq. No. 294

Contig ID 171_11.R1040

5'-most EST uC-gmronoir004f06b1

Seq. No. 295

Contig ID 171_12.R1040

5'-most EST LIB3093-004-Q1-K1-D4

Method BLASTX
NCBI GI g2780192
BLAST score 338
E value 2.0e-41
Match length 178
% identity 48

NCBI Description (AJ222713) unnamed protein product [Arabidopsis thaliana]

Seq. No. 296

Contig ID 171_13.R1040 5'-most EST kmv700739214.h1

Seq. No. 297

Contig ID 171_14.R1040

5'-most EST LIB3049-016-Q1-E1-A4

Method BLASTX
NCBI GI g4140398
BLAST score 375
E value 6.0e-36
Match length 105
% identity 63

NCBI Description (AF081794) sterol-C5(6)-desaturase [Nicotiana tabacum]

Seq. No. 298

Contig ID 171 15.R1040

5'-most EST jC-gmro02910059f03d1

Method BLASTX
NCBI GI g3152559
BLAST score 489



E value 1.0e-48 Match length 333 % identity 40

NCBI Description (AC002986) Similarity to A. thaliana gene product

F21M12.20, gb_AC000132. EST gb_Z25651 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 299

Contig ID 171_16.R1040 5'-most EST zvj700605110.h2

Method BLASTX
NCBI GI g3800878
BLAST score 664
E value 4.0e-86
Match length 201
% identity 63

NCBI Description (AF096281) threonine dehydratase/deaminase [Arabidopsis

thaliana]

Seq. No. 300

Contig ID 171_17.R1040

5'-most EST LIB3087-010-Q1-K1-A3

Method BLASTX
NCBI GI g1432083
BLAST score 349
E value 7.0e-51
Match length 155
% identity 70

NCBI Description (U60981) homolog to Skplp, an evolutionarily conserved

kinetochore protein in budding yeast [Arabidopsis thaliana] >gi_3068807 (AF059294) Skp1 homolog [Arabidopsis thaliana]

>gi 3719209 (U97020) UIP1 [Arabidopsis thaliana]

Seq. No. 301

Contig ID 171_18.R1040

5'-most EST uC-gmropic010a05b1

Seq. No. 302

Contig ID 171_19.R1040 5'-most EST wvk700683133.h1

Method BLASTN
NCBI GI g2150129
BLAST score 104
E value 5.0e-51
Match length 344
% identity 83

NCBI Description Arabidopsis thaliana cytoplasmic ribosomal protein S15a

mRNA, complete cds

Seq. No. 303

Contig ID 171_20.R1040

5'-most EST jC-gmst02400027h04d1

Method BLASTX
NCBI GI g3800878
BLAST score 376
E value 1.0e-35
Match length 136



% identity

(AF096281) threonine dehydratase/deaminase [Arabidopsis NCBI Description thaliana]

Seq. No.

304

Contig ID

171 21.R1040

5'-most EST

LIB3040-043-Q1-E1-A11

Method NCBI GI BLASTX q1173218

BLAST score E value

382 2.0e-41

Match length

110

% identity

90

NCBI Description

40S RIBOSOMAL PROTEIN S15A >gi 440824 (L27461) ribosomal protein S15 [Arabidopsis thaliana] >gi 2150130 (AF001412) cytoplasmic ribosomal protein S15a [Arabidopsis thaliana]

Seq. No.

Contig ID

171 22.R1040

5'-most EST

uC-gmropic104d08b1

Method NCBI GI BLASTX

BLAST score

q2760606 302

E value

1.0e-68

Match length

165 85

% identity NCBI Description

(AB001568) phospholipid hydroperoxide glutathione

peroxidase-like protein [Arabidopsis thaliana] >gi 3004869

(AF030132) glutathione peroxidase; ATGP1 [Arabidopsis

thaliana] >gi 4539451 emb CAB39931.1 (AL049500) phospholipid hydroperoxide glutathione peroxidase

[Arabidopsis thaliana]

Seq. No.

306

Contig ID

171_23.R1040

5'-most EST

jC-gmle01810010g10a1

Method NCBI GI BLASTN g20740 258

BLAST score

1.0e-142

E value Match length

% identity

813 84

NCBI Description Pisum sativum mRNA for P protein, a part of glycine

cleavage complex

Seq. No.

307

Contig ID 5'-most EST 171 24.R1040 vzy700754515.h1

Seq. No.

Method

308

Contig ID

171 25.R1040

5'-most EST

LIB3050-018-Q1-E1-C2 BLASTX

NCBI GI

q480618

BLAST score

629

E value Match length 2.0e-73

121



% identity 65
NCBI Description ATAF1 protein - Arabidopsis thaliana (fragment)
>gi_1345506_emb_CAA52771_ (X74755) ATAF1 [Arabidopsis thaliana]

Seq. No. 309

Contig ID 171_26.R1040

5'-most EST LIB3072-021-Q1-E1-C1

Method BLASTX
NCBI GI g3551838
BLAST score 331
E value 4.0e-49
Match length 151
% identity 75

NCBI Description (AF070967) SKP1-like protein [Nicotiana clevelandii]

Seq. No. 310

Contig ID 171_27.R1040

5'-most EST LIB3138-085-P1-N1-A12

Method BLASTN
NCBI GI g210811
BLAST score 163
E value 4.0e-86
Match length 704
% identity 85

NCBI Description Bean pod mottle virus coat protein gene, complete cds,

complete middle component (M) RNA

Seq. No. 311

Contig ID 171_28.R1040 5'-most EST awf700838388.h1

Method BLASTX
NCBI GI g480618
BLAST score 423
E value 6.0e-41
Match length 233
% identity 48

NCBI Description ATAF1 protein - Arabidopsis thaliana (fragment)

>gi 1345506_emb_CAA52771_ (X74755) ATAF1 [Arabidopsis

thaliana]

Seq. No. 312

Contig ID 171 29.R1040

5'-most EST LIB3107-010-Q1-K1-B8

Method BLASTX
NCBI GI g1350954
BLAST score 148
E value 5.0e-23
Match length 129
% identity 62

NCBI Description 40S RIBOSOMAL PROTEIN S15A >gi_422493_pir__S33498 ribosomal

protein DS15a - fruit fly (Drosophila melanogaster)
>gi_311606_emb_CAA79771_ (Z21673) ribosomal protein 15a

(40S subunit) [Drosophila melanogaster]

Seq. No. 313

Contig ID 171_30.R1040



```
LIB3049-047-Q1-E1-H10
5'-most EST
                   BLASTN
Method
                   q296357
NCBI GI
                   67
BLAST score
                   8.0e-29
E value
Match length
                   403
 % identity
*NCBI Description C.sinensis mRNA csa for salt-associated protein
 Seq. No.
                   171 31.R1040
 Contig ID
                   LIB3107-049-Q1-K1-D4
 5'-most EST
                   BLASTN
 Method
                   g940048
 NCBI GI
                   79
 BLAST score
                   5.0e-36
 E value
 Match length
                   127
                    91
 % identity
NCBI Description Phaseolus vulgaris clone XZT-205 mRNA, partial cds
                    315
 Seq. No.
                    171 32.R1040
 Contig ID
                    uC-gmrominsoy304f08b1
 5'-most EST
                    BLASTX
 Method
                    q4097880
 NCBI GI
                    2146
 BLAST score
                    0.0e + 00
 E value
                    1521
 Match length
 % identity
                   (U70866) polyprotein [Bean pod mottle virus]
 NCBI Description
                    316
 Seq. No.
                    172 1.R1040
 Contig ID
                    fC-gmf1700905758_{FL}
 5'-most EST
                    BLASTX
 Method
                    g2129855
 NCBI GI
 BLAST score
                    1745
                    0.0e+00
 E value
 Match length
                    372
                    87
 % identity
                    mitogen-activated protein kinase MMK2 (EC 2.7.1.-) -
 NCBI Description
                    alfalfa
                    317
 Seq. No..
                    172 2.R1040
 Contig ID
                    g5127047
 5'-most EST
                    BLASTN
 Method
                    g1204128
 NCBI GI
                    272
 BLAST score
                    1.0e-151
 E value
                    456
 Match length
                    90
  % identity
 NCBI Description M.sativa MMK2 mRNA for protein kinase
```

318

172 3.R1040

uC-gmronoir049d11b1

Seq. No.

Contig ID

5'-most EST



Method BLASTN
NCBI GI g1204128
BLAST score 169
E value 8.0e-90
Match length 325
% identity 88

NCBI Description M.sativa MMK2 mRNA for protein kinase

Seq. No. Contig ID

319 173 1.R1040

5'-most EST

fC-gmf1700905635 FL

Method BLASTX
NCBI GI g2246376
BLAST score 418
E value 1.0e-40
Match length 164
% identity 65

NCBI Description (Z86093) b-Zip DNA binding protein [Arabidopsis thaliana]

Seq. No.

320

Contig ID

174 1.R1040

5'-most EST

fC-qmf1700905531 FL

Method BLASTN
NCBI GI g2739045
BLAST score 1015
E value 0.0e+00
Match length 1015
% identity 100

NCBI Description Glycine max polyphosphoinositide binding protein Ssh2p

(SSH2) mRNA, complete cds

Seq. No.

321

Contig ID 5'-most EST

175_1.R1040 zhf700957320.h1

Method BLASTX
NCBI GI g1352660
BLAST score 219
E value 4.0e-17
Match length 178
% identity 31

NCBI Description COP-COATED VESICLE MEMBRANE PROTEIN P24 PRECURSOR

>gi_924850 (U26264) CHOp24 [Cricetulus griseus]

Seq. No.

322

Contig ID

176 1.R1040

5'-most EST

jC-gmro02800028f01a1

Method BLASTX
NCBI GI g3024013
BLAST score 761
E value 1.0e-80
Match length 304
% identity 51

NCBI Description EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT

(EIF-2-ALPHA) >gi_2706460_emb_CAA15918.1_ (AL021046)

translational initiation factor 2 alpha

[Schizosaccharomyces pombe]

% identity

60



323

Seq. No.

```
Contig ID
                   176 2.R1040
5'-most EST
                   jC-gmst02400020dc03d1
Method
                   BLASTN
NCBI GI
                   g2264313
BLAST score
                   39
E value
                   2.0e-12
                   47
Match length
                   96
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MOP10, complete sequence [Arabidopsis thaliana]
                   324
Seq. No.
                  177 1.R1040
Contig ID
5'-most EST
                   k11701213972.h1
                  BLASTX
Method
NCBI GI
                   g3080435
BLAST score
                   1657
E value
                   0.0e + 00
Match length
                  412
% identity
                  (AL022605) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                   178 1.R1040
5'-most EST
                   fC-gmf1700904235 FL
Method
                   BLASTX
                  g584882
NCBI GI
BLAST score
                   781
E value
                   3.0e-83
Match length
                   240
% identity
                   40
                  CYCLOARTENOL SYNTHASE (2,3-EPOXYSQUALENE--CYCLOARTENOL
NCBI Description
                  CYCLASE) >gi 452446 (U02555) cycloartenol synthase;
                   (S)-2,3-epoxysqualene mutase [Arabidopsis thaliana]
                   326
Seq. No.
                   179 1.R1040
Contig ID
5'-most EST
                   fC-gmf1700903662 FL
Method
                   BLASTX
NCBI GI
                   g2706515
BLAST score
                   1080
                   1.0e-118
E value
Match length
                  309
% identity
                   66
                  (Y12689) isoflavone reductase-like protein [Citrus x
NCBI Description
                  paradisi]
Seq. No.
                   327
Contig ID
                   180 1.R1040
5'-most EST
                   fC-gmf1700902268 FL
Method
                  BLASTX
NCBI GI
                   g1430846
BLAST score
                   665
E value
                   2.0e-69
Match length
                  220
```



```
NCBI Description (X99210) myb-related transcription factor [Lycopersicon esculentum]
```

Contig ID 180_2.R1040 5'-most EST ncj700977108.h1

Seq. No. 329

Contig ID 180 3.R1040 5'-most EST leu701150112.h1

Method BLASTX
NCBI GI g2827545
BLAST score 379
E value 1.0e-36
Match length 69
% identity 94

NCBI Description (AL021635) myb-like protein [Arabidopsis thaliana]

Seq. No. 330

Contig ID 181_1.R1040

5'-most EST fC-gmst700890686_FL

Method BLASTX
NCBI GI g3256068
BLAST score 325
E value 1.0e-29
Match length 135
% identity 50

NCBI Description (Y14068) Heat Shock Factor 3 [Arabidopsis thaliana]

Seq. No. 331

Contig ID 182_1.R1040

5'-most EST fC-gmro700844279_FL

Method BLASTX
NCBI GI g2687724
BLAST score 1622
E value 0.0e+00
Match length 318
% identity 100

NCBI Description (AJ003245) NADPH:isoflavone reductase [Glycine max]

Seq. No.

Contig ID 182_2.R1040 5'-most EST fua701040175.h1

332

Method BLASTN
NCBI GI g2687723
BLAST score 544
E value 0.0e+00
Match length 964
% identity 89

NCBI Description Glycine max mRNA for putative NADPH:isoflavone reductase

Seq. No. 333

Contig ID 182_3.R1040 5'-most EST g5057872 Method BLASTN NCBI GI g2687723

BLAST score 234



```
1.0e-129
E value
                  371
Match length
                  93
% identity
                  Glycine max mRNA for putative NADPH:isoflavone reductase
NCBI Description
Seq. No.
                  183 1.R1040
Contig ID
                  fC-gmst700888747_FL
5'-most EST
                  BLASTX
Method
                  g544015
NCBI GI
                  566
BLAST score
                   4.0e-58
E value
                  122
Match length
                   80
% identity
                  ENDOCHITINASE PRECURSOR >gi_2118033_pir__S59947 chitinase
NCBI Description
                   (EC 3.2.1.14) Al precursor - garden pea
                  >gi_20687_emb_CAA45359_ (X63899) chitinase [Pisum sativum]
                   335
Seq. No.
                   183 2.R1040
Contig ID
                   LIB3107-056-Q1-K1-F3
5'-most EST
                   BLASTX
Method
                   q544015
NCBI GI
                   1304
BLAST score
                   1.0e-144
E value
                   319
Match length
                   73
% identity
                   ENDOCHITINASE PRECURSOR >gi_2118033_pir__S59947 chitinase
NCBI Description
                   (EC 3.2.1.14) Al precursor - garden pea
                   >gi_20687_emb_CAA45359_ (X63899) chitinase [Pisum sativum]
Seq. No.
                   184 1.R1040
Contig ID
                   uC-gmrominsoy188g06b1
5'-most EST
                   BLASTX
Method
                   g2924777
NCBI GI
                   295
BLAST score
                   3.0e-26
E value
                   265
Match length
% identity
                   (AC002334) putative receptor protein kinase [Arabidopsis
NCBI Description
                   thaliana]
                   337
Seq. No.
                   185 1.R1040
 Contig ID
                   fC-gmst700889963 FL
 5'-most EST
                   BLASTX
Method
                   g1651197
NCBI GI
                   416
 BLAST score
                   1.0e-40
 E value
                   135
 Match length
                   56
 % identity
                   (D88156) tropinone reductase-I [Hyoscyamus niger]
 NCBI Description
```

Contig ID 186_1.R1040

5'-most EST fC-gmse700675825_FL



Contig ID 188_1.R1040

5'-most EST fC-gmst700889529_FL

Method BLASTX
NCBI GI g1518540
BLAST score 780
E value 3.0e-83
Match length 180
% identity 86

NCBI Description (U53418) UDP-glucose dehydrogenase [Glycine max]

Seq. No. 340

Contig ID 188_2.R1040 5'-most EST zhf700964379.h1

Method BLASTX
NCBI GI g1518540
BLAST score 164
E value 1.0e-11
Match length 34
% identity 97

NCBI Description (U53418) UDP-glucose dehydrogenase [Glycine max]

Seq. No. 341

Contig ID 189_1.R1040

5'-most EST fC-gmst700888192_FL

Method BLASTX
NCBI GI g1827893
BLAST score 314
E value 1.0e-28
Match length 93
% identity 61

NCBI Description Maize Nonspecific Lipid Transfer Protein Complexed With Palmitate >gi_1827894_pdb_1MZL_ Maize Nonspecific Lipid

Transfer Protein >gi_2194092_pdb_1AFH_ Lipid Transfer

Protein From Maize Seedlings, Nmr, 15 Structures

Seq. No. 342

Contig ID 190_1.R1040

5'-most EST fC-gmfl700902987_FL

Method BLASTN
NCBI GI g2627180
BLAST score 276
E value 1.0e-153
Match length 468
% identity 90

NCBI Description Pisum sativum mRNA for cycloartenol synthase, complete cds

Seq. No. 343

Contig ID 191_1.R1040

5'-most EST LIB3139-030-P1-N1-A6

Seq. No. 344

Contig ID 192_1.R1040

5'-most EST fC-gmro700845184_FL

Method BLASTX NCBI GI g2465010



```
BLAST score
                   3.0e-41
E value
                   141
Match length
                   63
% identity
NCBI Description (AJ001446) acyl carrier protein [Fragaria vesca]
Seq. No.
                   192 2.R1040
Contig ID
                   fua701038142.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2465010
                   382
BLAST score
                   2.0e-36
E value
                   139
Match length
% identity
                   (AJ001446) acyl carrier protein [Fragaria vesca]
NCBI Description
Seq. No.
                   193 1.R1040
Contig ID
                   fC-qmst700790793 FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1708236
                   1861
BLAST score
                   0.0e + 00
E value
                    446
Match length
                    77
% identity
                   HYDROXYMETHYLGLUTARYL-COA SYNTHASE (HMG-COA SYNTHASE)
NCBI Description
                    (3-HYDROXY-3-METHYLGLUTARYL COENZYME A SYNTHASE)
                   >gi_2129617_pir__JC4567 hydroxymethylglutaryl-CoA synthase
                    (EC 4.1.3.5) - Arabidopsis thaliana
                   >gi_1143390_emb_CAA58763_ (X83882)
hydroxymethylglutaryl-CoA synthase [Arabidopsis thaliana]
                    >gi_1586548_prf__2204245A hydroxy methylglutaryl CoA
                    synthase [Arabidopsis thaliana]
                    347
 Seq. No.
                    194 1.R1040
 Contig ID
                    fC-zmf1700337153 FL
 5'-most EST
                    BLASTX
 Method
                    g1155255
 NCBI GI
 BLAST score
                    323
                    1.0e-29
 E value
                    130
 Match length
                    48
 % identity
                   (U39228) beta-glucosidase [Prunus avium]
 NCBI Description
                    348
 Seq. No.
                    195_1.R1040
 Contig ID
                    fC-zmf1700354918_FL
 5'-most EST
                    BLASTX
 Method
                    g548493
 NCBI GI
 BLAST score
                    916
                    7.0e-99
 E value
                    289
 Match length
```

(GALACTURAN 1,4-ALPHA-GALACTURONIDASE)

EXOPOLYGALACTURONASE PRECURSOR (EXOPG) (PECTINASE)

59

% identity

NCBI Description



>gi_629854_pir__S30067 polygalacturonase - maize >gi 288612 emb CAA47052 (X66422) polygalacturonase [Zea mays]

349 Seq. No.

196_1.R1040 Contig ID

fC-zmf1700354921_FL 5'-most EST

BLASTN Method q429006 NCBI GI 39 BLAST score 3.0e-12 E value 95 Match length 85 % identity

NCBI Description Rice mRNA for MCM3 (gene name SS300), partial cds

Seq. No.

197 1.R1040 Contig ID

fC-zmf1700357449 FL 5'-most EST

BLASTX Method q3402722 NCBI GI 1200 BLAST score 1.0e-132 E value 302 Match length

75 % identity

NCBI Description (AC004261) CPDK-related protein [Arabidopsis thaliana]

351 Seq. No.

198 1.R1040 Contig ID

fC-zmf1700379641_FL 5'-most EST

BLASTX Method q3342242 NCBI GI 1840 BLAST score 0.0e + 00E value 463 Match length 72 % identity

(AF030421) cell wall invertase; beta-fructofuranosidase; NCBI Description

fructosidase [Triticum aestivum]

352 Seq. No.

199_1.R1040 Contig ID

fC-zmf1700379870 FL 5'-most EST

Method BLASTX q3057120 NCBI GI 1314 BLAST score 1.0e-145 E value 292 Match length 82 % identity

(AF023159) starch synthase DULL1 [Zea mays] NCBI Description

353 Seq. No.

200 1.R1040 Contig ID

fC-zmro700829959_FL 5'-most EST

BLASTX Method q4154352 NCBI GI 392 BLAST score 1.0e-37 E value Match length 241



% identity 39
NCBI Description (AF110333) PrMC3 [Pinus radiata]

Seq. No. 354

Contig ID 202 1.R1040

5'-most EST fC-zmro700831955_FL

Method BLASTX
NCBI GI g3850630
BLAST score 700
E value 1.0e-73
Match length 199
% identity 62

NCBI Description (AJ012581) cytochrome P450 [Cicer arietinum]

Seq. No. 355

Contig ID 202_2.R1040

5'-most EST fC-zmro700831544_FL

Method BLASTX
NCBI GI g3850630
BLAST score 234
E value 1.0e-36
Match length 119
% identity 54

NCBI Description (AJ012581) cytochrome P450 [Cicer arietinum]

Seq. No. 356

Contig ID 202 3.R1040 5'-most EST 94396629 Method BLASTX NCBI GI 94200044 BLAST score 1306 E value 1.0e-144 Match length 382

Match length 382 % identity 63

NCBI Description (AB022732) cytochrome P450 [Glycyrrhiza echinata]

Seq. No. 357

Contig ID 203 1.R1040

5'-most EST fC-zmro700832243 FL

Method BLASTX
NCBI GI g3335378
BLAST score 614
E value 2.0e-63
Match length 234
% identity 54

NCBI Description (AC003028) Myb-related transcription activator [Arabidopsis

thaliana]

Seq. No. 358

Contig ID 204 1.R1040

5'-most EST fC-gmle700870704h2

Method BLASTX
NCBI GI g2129770
BLAST score 1151
E value 1.0e-126
Match length 291
% identity 71



NCBI Description xyloglucan endotransglycosylase-related protein XTR-2 -Arabidopsis thaliana >gi 1244756 (U43487) xyloglucan

endotransglycosylase-related protein [Arabidopsis thaliana]

>gi_2154611_dbj_BAA20290_ (D63510) endoxyloglucan transferase related protein [Arabidopsis thaliana]

Seq. No.

359

Contig ID

204_2.R1040

5'-most EST

fC-qmse7000756651rl BLASTX

Method NCBI GI

q2129770

BLAST score E value

517 2.0e-52

Match length % identity

145

NCBI Description

66 xyloglucan endotransglycosylase-related protein XTR-2 -Arabidopsis thaliana >gi 1244756 (U43487) xyloqlucan

endotransglycosylase-related protein [Arabidopsis thaliana]

>gi 2154611 dbj BAA20290 (D63510) endoxyloglucan transferase related protein [Arabidopsis thaliana]

Seq. No.

Contig ID

204 3.R1040

5'-most EST

LIB3051-115-Q1-K1-D8

Method NCBI GI BLASTX g2129770

BLAST score E value

415 1.0e-40

Match length

98 73

% identity NCBI Description

xyloglucan endotransglycosylase-related protein XTR-2 -Arabidopsis thaliana >gi 1244756 (U43487) xyloglucan

endotransglycosylase-related protein [Arabidopsis thaliana]

>gi_2154611_dbj_BAA20290_ (D63510) endoxyloglucan transferase related protein [Arabidopsis thaliana]

Seq. No.

361

Contig ID 5'-most EST 204 4.R1040 LIB3107-038-Q1-K1-C9

Method

BLASTX

NCBI GI

g2129770

BLAST score

109

E value

8.0e-13

Match length % identity

52 70

NCBI Description

xyloglucan endotransglycosylase-related protein XTR-2 -Arabidopsis thaliana >qi 1244756 (U43487) xyloqlucan

endotransglycosylase-related protein [Arabidopsis thaliana]

>gi 2154611 dbj BAA20290 (D63510) endoxyloglucan transferase related protein [Arabidopsis thaliana]

Seq. No.

362

Contig ID 5'-most EST 204 5.R1040 dpv701102314.h1

Method NCBI GI

BLASTX q2129770

BLAST score

177



E value 7.0e-13 Match length 45 % identity 71

NCBI Description xyloglucan endotransglycosylase-related protein XTR-2 - Arabidopsis thaliana >gi 1244756 (U43487) xyloglucan

endotransglycosylase-related protein [Arabidopsis thaliana]

>gi_2154611_dbj_BAA20290_ (D63510) endoxyloglucan transferase related protein [Arabidopsis thaliana]

Seq. No. 363

Contig ID 205 1.R1040

5'-most EST fC-zmro700833257_FL

Method BLASTX
NCBI GI g2494329
BLAST score 378
E value 8.0e-36
Match length 310
% identity 34

NCBI Description GLUCAN 1,3-BETA-GLUCOSIDASE PRECURSOR

(EXO-1,3-BETA-GLUCANASE) >gi_1150694 emb_CAA86952_ (Z46872) exo-1,3-beta-glucanase/1,3-beta-D-glucan glucanohydrolase

[Yarrowia lipolytica]

Seq. No. 364

Contig ID 205_2.R1040 5'-most EST zsg701127354.h1

Method BLASTX
NCBI GI g1050956
BLAST score 295
E value 4.0e-26
Match length 157
% identity 46

NCBI Description (U26160) 43 kDa secreted glycoprotein precursor

[Paracoccidioides brasiliensis] >gi_1588394_prf__2208385A

glycoprotein gp43 [Paracoccidioides brasiliensis]

Seq. No. 365

Contig ID 206_1.R1040

Method BLASTX
NCBI GI g3128195
BLAST score 1524
E value 1.0e-170
Match length 361
% identity 83

NCBI Description (AC004521) putative phosphoribosyl pyrophosphate synthetase

[Arabidopsis thaliana] >gi_3341673 (AC003672) putative phosphoribosyl pyrophosphate synthetase [Arabidopsis

thaliana]

Seq. No. 366

Contig ID 207_1.R1040

5'-most EST fC-zmro700833923 FL

Method BLASTX
NCBI GI g3540200
BLAST score 237
E value 2.0e-19



Match length 186 % identity 35

NCBI Description (AC004260) Similar to TINY [Arabidopsis thaliana]

Seq. No. Contig ID

367

5'-most EST

208_1.R1040 fC-zmro700834004 FL

Method NCBI GI BLASTX g2979553

BLAST score E value Match length 511 1.0e-51 166

% identity NCBI Description 59 (AC003680) hypothetical protein [Arabidopsis thaliana]

Seq. No.

368

Contig ID

210 1.R1040

5'-most EST

fC-gmro700845739 FL

Seq. No.

369

Contig ID

210 3.R1040

5'-most EST

LIB3170-026-Q1-J1-F6

Seq. No.

370

Contig ID

210_4.R1040

5'-most EST

LIB3107-065-Q1-K1-B11

Seq. No.

371

Contig ID

211 1.R1040

5'-most EST

fC-gmro700849012_FL

Method NCBI GI BLASTX g3643607

BLAST score

193

E value

1.0e-14

Match length % identity

45 78

NCBI Description

(AC005395) unknown protein [Arabidopsis thaliana]

Seq. No.

372

Contig ID

212_1.R1040

5'-most EST

fC-gmst700791347 FL

Method NCBI GI BLASTX g3643601

BLAST score

252

E value Match length 3.0e-21

% identity

96 54

NCBI Description

(AC005395) hypothetical protein [Arabidopsis thaliana]

Seq. No.

373

Contig ID 5'-most EST

213_3.R1040 bth700845603.h1

Method NCBI GI BLASTN q2804153

BLAST score 304 E value 1.0e-170

E value Match length

693



% identity

NCBI Description Lupinus albus mRNA for aquaporin, partial

Seq. No.

Contig ID

213 4.R1040

5'-most EST

LIB3073-005-Q1-K1-G8

Method NCBI GI BLASTX g1175012

BLAST score E value

1285 1.0e-142

Match length

283

% identity

86

NCBI Description

PLASMA MEMBRANE INTRINSIC PROTEIN 1C (TRANSMEMBRANE PROTEIN B) (TMP-B) >gi_396218_emb_CAA49155_ (X69294) transmembrane

protein TMP-B [Arabidopsis thaliana]

375 Seq. No.

Contig ID

213 5.R1040

5'-most EST Method

leu701152670.hl BLASTN

NCBI GI

g3037046

BLAST score E value

402 0.0e + 00

Match length % identity

656 91

NCBI Description

Phaseolus vulgaris 1-aminocyclopropane-1-carboxylic acid

oxidase mRNA, complete cds

Seq. No.

376

Contig ID

213 6.R1040

5'-most EST Method

LIB3040-044-Q1-E1-D4 BLASTN

NCBI GI BLAST score q2385377

E value

162 2.0e-85

Match length

517

% identity

85

NCBI Description Nicotiana tabacum mRNA for aquaporin

Seq. No.

Contig ID 5'-most EST 213 7.R1040 g1575728 FL

Method

BLASTN

NCBI GI

g1575728

BLAST score

638

E value

0.0e + 00

Match length

790

% identity

NCBI Description

95 Glycine max 14-3-3 related protein SGF14C mRNA, complete

cds

Seq. No.

378

Contig ID

213 8.R1040

5'-most EST

uC-gmropic107f05b1

Method

BLASTN

NCBI GI

g2149954

BLAST score

221



```
1.0e-120
E value
                   801
Match length
                   88
% identity
                  Phaseolus vulgaris putative aquaporin-1 (Mip-1) mRNA,
NCBI Description
                   complete cds
                   379
Seq. No.
                   214 1.R1040
Contig ID
                   LIB3170-004-Q1-K1-H9
5'-most EST
                   BLASTX
Method
                   q1149595
NCBI GI
                   1054
BLAST score
                   1.0e-115
E value
                   256
Match length
                   77
% identity
                   (Z49860) 1-acyl-sn-glycerol-3-phosphate acyltransferase
NCBI Description
                   [Brassica napus]
                   380
Seq. No.
                   214 2.R1040
Contig ID
                   jC-gmro02910070h01a1
5'-most EST
                   BLASTX
Method
                   g1149595
NCBI GI
                   200
BLAST score
E value
                   2.0e-15
                   63
Match length
                   60
% identity
                   (Z49860) 1-acyl-sn-glycerol-3-phosphate acyltransferase
NCBI Description
                    [Brassica napus]
                   381
Seq. No.
                   215 1.R1040
Contig ID
                   g56\overline{0}5705
5'-most EST
                   BLASTX
Method
                   q710626
NCBI GI
                   155
BLAST score
                    6.0e-10
E value
                    42
Match length
                    57
% identity
                    (D30719) ERD15 protein [Arabidopsis thaliana] >gi_3241941
NCBI Description
                    (AC004625) dehydration-induced protein ERD15 [Arabidopsis
                    thaliana] >gi_3894181 (AC005662) ERD15 protein [Arabidopsis
                    thaliana]
 Seq. No.
                    382
                    215 2.R1040
 Contig ID
                    a4291035
 5'-most EST
                    BLASTX
Method
NCBI GI
                    q710626
                    147
 BLAST score
                    5.0e-09
 E value
                    42
 Match length
                    57
 % identity
                    (D30719) ERD15 protein [Arabidopsis thaliana] >gi_3241941
 NCBI Description
                    (AC004625) dehydration-induced protein ERD15 [Arabidopsis
```

thaliana]

thaliana] >gi_3894181 (AC005662) ERD15 protein [Arabidopsis



```
Seq. No.
                   383
Contig ID
                  215 3.R1040
5'-most EST
                  uC-gmrominsoy276h04b1
Method
                  BLASTX
NCBI GI
                  g710626
BLAST score
                  155
E value
                  4.0e-10
Match length
                  42
% identity
                  57
                  (D30719) ERD15 protein [Arabidopsis thaliana] >gi 3241941
NCBI Description
                   (AC004625) dehydration-induced protein ERD15 [Arabidopsis
                  thaliana] >gi 3894181 (AC005662) ERD15 protein [Arabidopsis
                  thaliana]
                  384
Seq. No.
Contig ID
                  215 4.R1040
5'-most EST
                  uC-gmrominsoy049c05b1
Method
                  BLASTX
NCBI GI
                  g710626
BLAST score
                  155
E value
                  3.0e-10
Match length
                  42
% identity
                  57
                  (D30719) ERD15 protein [Arabidopsis thaliana] >gi 3241941
NCBI Description
                   (AC004625) dehydration-induced protein ERD15 [Arabidopsis
                  thaliana] >gi 3894181 (AC005662) ERD15 protein [Arabidopsis
                  thaliana]
                  385
Seq. No.
Contig ID
                  216 1.R1040
5'-most EST
                  leu701148094.h1
Method
                  BLASTX
NCBI GI
                  g3980406
BLAST score
                  346
E value
                  2.0e-32
Match length
                  103
% identity
                  57
                  (AC004561) putative tropinone reductase [Arabidopsis
NCBI Description
                  thalianal
                  386
Seq. No.
                  218 1.R1040
Contig ID
5'-most EST
                  g5057891
```

Contig ID 218_1.R1040
5'-most EST g5057891
Method BLASTX
NCBI GI g585973
BLAST score 346
E value 2.0e-32
Match length 89
% identity 76

NCBI Description FRUCTOKINASE >gi_626018_pir__S39997 fructokinase (EC 2.7.1.4) - potato >gi_297015_emb_CAA78283 (Z12823)

fructokinase [Solanum tuberosum] >gi 1095321 prf 2108342A

fructokinase [Solanum tuberosum]

Seq. No.

387

Contig ID

219 1.R1040

Contig ID

392

223 1.R1040



```
rca700997536.h1
5'-most EST
                  BLASTX
Method
                  g1653655
NCBI GI
BLAST score
                  364
E value
                  5.0e-34
Match length
                  163
% identity
                  44
NCBI Description
                  (D90915) ATP-dependent Clp protease proteolytic subunit
                   [Synechocystis sp.]
Seq. No.
                  388
                  220 1.R1040
Contig ID
5'-most EST
                   fC-qmro700844894 FL
Method
                  BLASTX
                  g4512651
NCBI GI
BLAST score
                  616
                  7.0e-64
E value
Match length
                  222
% identity
                  (AC007048) putative tyrosine transaminase [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  389
                  221 1.R1040
Contig ID
5'-most EST
                  jC-gmst02400054a06d1
                  BLASTX
Method
NCBI GI
                  g349379
BLAST score
                  337
E value
                  3.0e-31
Match length
                  127
                  63
% identity
                  (L22847) HAHB-1 [Helianthus annuus]
NCBI Description
                  390
Seq. No.
                  221 2.R1040
Contig ID
5'-most EST
                  gsv701054757.h1
Method
                  BLASTX
NCBI GI
                  q349379
BLAST score
                   295
                  2.0e-26
E value
Match length
                  129
                  57
% identity
NCBI Description
                  (L22847) HAHB-1 [Helianthus annuus]
Seq. No.
                  391
Contig ID
                   222 1.R1040
5'-most EST
                   fC-gmro700847730 FL
                  BLASTX
Method
NCBI GI
                  g4415914
BLAST score
                   367
E value
                  1.0e-34
Match length
                  198
% identity
                   37
                  (AC006282) unknown protein [Arabidopsis thaliana]
NCBI Description
```



5'-most EST fC-gmle7000741669a1 Method BLASTX g1169586 NCBI GI BLAST score 1610 E value 1.0e-180 338

Match length 89 % identity

NCBI Description FRUCTOSE-1, 6-BISPHOSPHATASE, CYTOSOLIC

(D-FRUCTOSE-1, 6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE) (CY-F1) >gi_542079_pir__S41287 fructose-bisphosphatase (EC 3.1.3.11) - potato >gi_440591_emb_CAA54265_ (X76946)

fructose-1, 6-bisphosphatase [Solanum tuberosum]

Seq. No. 393

223 2.R1040 Contig ID

5'-most EST fC-gmle7000741457a1

Method BLASTX NCBI GI g3913640 BLAST score 591 E value 4.0e-61 120 Match length % identity

FRUCTOSE-1, 6-BISPHOSPHATASE, CYTOSOLIC NCBI Description

(D-FRUCTOSE-1, 6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)

>gi 3041775 dbj BAA25422 (AB007193)

fructose-1,6-bisphosphatase [Oryza sativa]

394 Seq. No.

223 3.R1040 Contig ID 5'-most EST ncj700986460.h1

Method BLASTX q3913640 NCBI GI BLAST score 523 E value 3.0e-53 Match length 111 86 % identity

FRUCTOSE-1, 6-BISPHOSPHATASE, CYTOSOLIC NCBI Description

(D-FRUCTOSE-1, 6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)

>gi 3041775 dbj BAA25422 (AB007193)

fructose-1,6-bisphosphatase [Oryza sativa]

395 Seq. No.

Contig ID 223 4.R1040

fC-gmle700873234a1 5'-most EST

Method BLASTX NCBI GI g2494416 BLAST score 434 E value 2.0e-61 127 Match length % identity 92

NCBI Description FRUCTOSE-1,6-BISPHOSPHATASE, CYTOSOLIC

(D-FRUCTOSE-1, 6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE) >gi_166956 (M80597) fructose-1,6-bisphosphatase [Beta vulgaris] >gi_444324_prf__1906373A cytosolic fructose

bisphosphatase [Beta vulgaris]

Seq. No. 396



```
224 1.R1040
Contig ID
5'-most EST
                   fC-gmro700849195 FL
Method
                  BLASTX
NCBI GI
                   q4490335
BLAST score
                   267
E value
                   5.0e-23
Match length
                   153
% identity
                   44
NCBI Description
                   (AL035656) receptor kinase-like protein [Arabidopsis
                   thaliana]
                   397
Seq. No.
                   224 2.R1040
Contig ID
5'-most EST
                   fC-qmle700875134r1
                   398
Seq. No.
Contig ID
                   225 1.R1040
5'-most EST
                  LIB3092-012-Q1-K1-A12
                   399
Seq. No.
Contig ID
                   226 1.R1040
5'-most EST
                   fC-qmro700864101 FL
                   BLASTX
Method
NCBI GI
                   g2924777
BLAST score
                   187
E value
                   8.0e-14
Match length
                   107
% identity
                   43
NCBI Description
                  (AC002334) putative receptor protein kinase [Arabidopsis
                   thaliana]
                   400
Seq. No.
                   226 2.R1040
Contig ID
5'-most EST
                  LIB3139-100-P1-N1-A12
                   401
Seq. No.
Contig ID
                   227 1.R1040
5'-most EST
                   fC-gmro700865543 FL
Method
                   BLASTX
NCBI GI
                   q4455129
BLAST score
                   317
                   1.0e-28
E value
                   120
Match length
% identity
                   52
                  (AF127761) ribonucleoprotein ZRNP1 [Homo sapiens]
NCBI Description
                   402
Seq. No.
                   227 2.R1040
Contig ID
                   LIB3106~022-Q1-K1-E1
5'-most EST
Method
                   BLASTX
```

Method BLASTX
NCBI GI g4455129
BLAST score 201
E value 2.0e-16
Match length 80
% identity 61

NCBI Description (AF127761) ribonucleoprotein ZRNP1 [Homo sapiens]

% identity

60



```
Seq. No.
                  228 1.R1040
Contig ID
                  fC-gmro700868230_FL
5'-most EST
                  BLASTX
Method
                  q4544403
NCBI GI
                  720
BLAST score
                  5.0e-76
E value
                  206
Match length
                   62
% identity
NCBI Description (AC007047) putative glucan endo-1,3-beta-D-glucosidase
                  precursor [Arabidopsis thaliana]
                   404
Seq. No.
                   228 2.R1040
Contig ID
                  uC-gmrominsoy169b02b1
5'-most EST
                  BLASTX
Method
                   g3641838
NCBI GI
                   647
BLAST score
                   1.0e-132
E value
                   329
Match length
% identity
                  (AL023094) putative protein (fragment) [Arabidopsis
NCBI Description
                   thaliana]
                   405
Seq. No.
                   229 1.R1040
Contig ID
                   fC-gmle700743613f1
5'-most EST
                   BLASTN
Method
                   q435678
NCBI GI
BLAST score
                   132
                   9.0e-68
E value
                   323
Match length
% identity
                   L.esculentum Mill (cv. Rutgers) mRNA for ribosomal protein
NCBI Description
                   406
Seq. No.
                   229 2.R1040
Contig ID
                   gsv701055480.h1
 5'-most EST
                   BLASTN
Method
                   q435678
NCBI GI
BLAST score
                   120
E value
                   1.0e-60
                   323
Match length
 % identity
                   85
                   L.esculentum Mill (cv. Rutgers) mRNA for ribosomal protein
 NCBI Description
                   S25
                   407
 Seq. No.
                   229 3.R1040
 Contig ID
                   V4L-01-Q1-B1-E8
 5'-most EST
                   BLASTX
 Method
                   g3249084
 NCBI GI
                   1818
 BLAST score
                   0.0e + 00
 E value
                    567
 Match length
```



NCBI Description (AC004473) Similar to red-1 (related to thioredoxin) gene gb_X92750 from Mus musculus. ESTs gb_AA712687 and qb_Z37223 come from this gene [Arabidopsis thaliana]

Seq. No. 408

Contig ID 229 4.R1040 5'-most EST cfl700863563.h1

Method BLASTX
NCBI GI g4567232
BLAST score 239
E value 1.0e-30
Match length 83
% identity 85

NCBI Description (AC007119) putative 40S ribosomal protein S25 [Arabidopsis

thaliana]

Seq. No. 409

Contig ID 230_1.R1040

5'-most EST uC-gmronoir037a08b1

Method BLASTX
NCBI GI g1922964
BLAST score 609
E value 8.0e-63
Match length 171
% identity 73

NCBI Description (AC000106) Similar to Schizosaccharomyces CCAAT-binding

factor (gb U88525). EST gb_T04310 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 410

Contig ID 231_1.R1040

5'-most EST fC-gmro700566893r1

Method BLASTX
NCBI GI g4538929
BLAST score 672
E value 4.0e-70
Match length 362
% identity 42

NCBI Description (AL049483) putative nucleic acid binding protein

[Arabidopsis thaliana]

Seq. No. 411

Contig ID 231_2.R1040

5'-most EST jC-gmle01810016c05d1

Method BLASTX
NCBI GI g2342727
BLAST score 361
E value 3.0e-34
Match length 90
% identity 70

NCBI Description (AC002341) hypothetical protein [Arabidopsis thaliana]

Seq. No. 412

Contig ID 231_3.R1040 5'-most EST fua701043550.h1

Method BLASTX NCBI GI g4538929



BLAST score 1.0e-25 E value Match length 175 49 % identity

(AL049483) putative nucleic acid binding protein NCBI Description

[Arabidopsis thaliana]

Seq. No.

231 4.R1040 Contig ID

ASG32442DA-02-Q1-E1-F10 5'-most EST BLASTX Method

q4538929 NCBI GI 163 BLAST score 5.0e-11 E value 58 Match length 55 % identity

(AL049483) putative nucleic acid binding protein NCBI Description

[Arabidopsis thaliana]

Seq. No.

232 1.R1040 Contig ID

fC-zmro700830348 FL 5'-most EST

BLASTX Method q3901012 NCBI GI 1157 BLAST score 1.0e-127 E value 278 Match length 76

% identity (AJ130885) xyloglucan endotransglycosylase 1 [Fagus NCBI Description

sylvatica]

415 Seq. No.

232 2.R1040 Contig ID

uC-gmropic018b01b1 5'-most EST

BLASTX Method q3901012 NCBI GI 313 BLAST score 9.0e-29 E value 75 Match length % identity

(AJ130885) xyloglucan endotransglycosylase 1 [Fagus NCBI Description

sylvatica]

Seq. No. 416

232 3.R1040 Contig ID gsv701049278.h1 5'-most EST

BLASTX Method NCBI GI g2129771 BLAST score 334 5.0e-31 E value 164 Match length % identity 68 NCBI Description

xyloglucan endotransglycosylase-related protein XTR-6 -Arabidopsis thaliana >gi_1244758 (U43488) xyloglucan

endotransglycosylase-related protein [Arabidopsis thaliana] >gi_4539299 emb CAB39602.1 (AL049480) xyloglucan endo-1, 4-beta-D-glucanase (XTR-6) [Arabidopsis thaliana]



```
417
Seq. No.
                   232 4.R1040
Contig ID
                   zzp700835657.h1
5'-most EST
                  BLASTX
Method
                   q3901012
NCBI GI
                   645
BLAST score
E value
                   2.0e-67
                   146
Match length
                   79
% identity
                   (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
NCBI Description
                   sylvatica]
                   418
Seq. No.
                   232 6.R1040
Contig ID
                   LIB3139-037-P1-N1-B7
5'-most EST
                   BLASTX
Method
                   g3901012
NCBI GI
BLAST score
                   169
                   5.0e-12
E value
                   34
Match length
% identity
                   82
                   (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
NCBI Description
                   sylvatica]
                   419
Seq. No.
                   233 1.R1040
Contig ID
                   fC-zmro700832181 FL
5'-most EST
                   BLASTX
Method
                   g266685
NCBI GI
                   969
BLAST score
                   1.0e-105
E value
                   439
Match length
% identity
                   DIHYDROLIPOAMIDE ACETYLTRANSFERASE COMPONENT OF PYRUVATE
NCBI Description
                   DEHYDROGENASE COMPLEX (E2) (PDC-E2) (70 KD MITOCHONDRIAL
                   AUTOANTIGEN OF PRIMARY BILIARY CIRRHOSIS) (PBC)
                   >gi_111580_pir__S21766 dihydrolipoamide S-acetyltransferase
                    (EC_{2.3.1.12}) - rat
                    420
 Seq. No.
                    235 1.R1040
 Contig ID
 5'-most EST
                    fC-zmf1700381954 FL
 Method
                    BLASTN
                    g22485
 NCBI GI
 BLAST score
                    798
                    0.0e + 00
 E value
                    1075
 Match length
 % identity
                    98
                   Maize mRNA for sucrose synthase (EC 2.4.1.13)
 NCBI Description
                    421
 Seq. No.
 Contig ID
                    236 1.R1040
                    fC-zmf1700550326_FL
 5'-most EST
```

BLASTX

300

g4426627

Method

NCBI GI

BLAST score



```
3.0e-27
E value
                  59
Match length
% identity
                   (AF099969) sterol-C5(6)-desaturase homolog [Nicotiana
NCBI Description
                  tabacum]
                  422
Seq. No.
                  238 1.R1040
Contig ID
                  LIB3138-046-Q1-N1-E1
5'-most EST
                  BLASTX
Method
                  q1834387
NCBI GI
BLAST score
                  1830
                   0.0e + 00
E value
                   389
Match length
                   88
% identity
NCBI Description (Y11022) chalcone synthase [Betula pendula]
Seq. No.
                   239 1.R1040
Contig ID
                   fC-gmle700557774_FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3786006
                   589
BLAST score
                   2.0e-60
E value
                   177
Match length
                   67
% identity
NCBI Description (AC005499) hypothetical protein [Arabidopsis thaliana]
                   424
Seq. No.
Contig ID
                   240 1.R1040
                   fC-gmle700555604a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3242708
BLAST score
                   699
                   2.0e-73
E value
Match length
                   295
% identity
                   48
                   (AC003040) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
                   425
Seq. No.
                   240 2.R1040
Contig ID
                   LIB3139-045-P1-N1-C3
5'-most EST
                   BLASTX
Method
                   g3242708
NCBI GI
BLAST score
                   258
                   4.0e-22
E value
                   121
Match length
                   45
% identity
                   (AC003040) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
```

Seq. No. 426

Contig ID 241_1.R1040

5'-most EST fC-gmle700560462_FL

Method BLASTN NCBI GI g2463568

% identity

NCBI Description

80



```
BLAST score
                   5.0e-58
E value
Match length
                  203
                   89
% identity
NCBI Description Glycine max mRNA for squalene synthase, complete cds
                   427
Seq. No.
                   241 2.R1040
Contig ID
                   jC-gmle01810087d08d1
5'-most EST
                   BLASTN
Method
                   q2463568
NCBI GI
                   48
BLAST score
                   3.0e-18
E value
Match length
                   124
% identity
                   84
                   Glycine max mRNA for squalene synthase, complete cds
NCBI Description
Seq. No.
                   242 1.R1040
Contig ID
                   jC-gmro02910002d12a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3805849
                   4092
BLAST score
                   0.0e+00
E value
                   897
Match length
                   86
% identity
                   (AL031986) cytoplasmatic aconitate hydratase (citrate
NCBI Description
                   hydro-lyase) (aconitase) (EC 4.2.1.3) [Arabidopsis thaliana]
Seq. No.
                   243 1.R1040
Contig ID
                   epx701108931.hl
5'-most EST
                   BLASTX
Method
                   q2499488
NCBI GI
                   2683
BLAST score
                   0.0e + 00
E value
                   610
Match length
% identity
                   PYROPHOSPHATE--FRUCTOSE 6-PHOSPHATE 1-PHOSPHOTRANSFERASE
NCBI Description
                   ALPHA SUBUNIT (PFP) (6-PHOSPHOFRUCTOKINASE (PYROPHOSPHATE))
                   (PYROPHOSPHATE-DEPENDENT 6-PHOSPHOFRUCTOSE-1-KINASE)
                   (PPI-PFK) >gi_483547_emb_CAA83682_ (Z32849)
                   pyrophosphate-dependent phosphofructokinase alpha subunit
                   [Ricinus communis]
                   430
 Seq. No.
                   243 2.R1040
 Contig ID
                   g5688064
 5'-most EST
Method
                   BLASTX
 NCBI GI
                   g3790102
 BLAST score
                   432
                   2.0e-42
 E value
 Match length
                   102
```

alpha subunit [Citrus X paradisi]

(AF095521) pyrophosphate-dependent phosphofructokinase



Seq. No. 431

Contig ID 243_4.R1040 5'-most EST jex700904315.h1

Method BLASTN
NCBI GI g3790101
BLAST score 44
E value 1.0e-15
Match length 80
% identity 89

NCBI Description Citrus X paradisi pyrophosphate-dependent

phosphofructokinase alpha subunit (PPi-PFKa) mRNA, complete

cds

Seq. No. 432

Contig ID 244 1.R1040 5'-most EST ssr700556101.h1

Method BLASTX
NCBI GI g3915599
BLAST score 321
E value 3.0e-29
Match length 160
% identity 47

NCBI Description FLORAL HOMEOTIC PROTEIN AGL3 >gi_1361993_pir__S57793

MADS-box protein AGL3 - Arabidopsis thaliana >gi_1737495

(U81369) MADS box protein [Arabidopsis thaliana]

>gi_4406762_gb_AAD20073_ (AC006836) MADS box protein AGL3

[Arabidopsis thaliana]

Seq. No. 433

Contig ID 244_2.R1040 5'-most EST hyd700731174.h1

Seq. No. 434

Contig ID 245_1.R1040

5'-most EST fC-gmse700606016 FL

Method BLASTX
NCBI GI g2129788
BLAST score 161
E value 8.0e-11
Match length 38
% identity 84

NCBI Description biotin carboxyl carrier protein (clone BP7) - rape

(fragment) >gi_1070010_emb_CAA62266_ (X90732) Biotin

carboxyl carrier protein [Brassica napus]

>gi 1589045_prf__2210244F Ac-CoA carboxylase:ISOTYPE=bp7

[Brassica napus]

Seq. No. 435

Contig ID 246_1.R1040

5'-most EST LIB3051-115-Q1-K1-C12

Method BLASTX
NCBI GI g3335372
BLAST score 319
E value 7.0e-29
Match length 270
% identity 34

NCBI Description (AC003028) putative SRG1 protein [Arabidopsis thaliana]

NCBI GI

E value

BLAST score



```
Seq. No.
                   436
Contig ID
                  247 1.R1040
5'-most EST
                   fC-gmle700553710 FL
Method
                  BLASTX
NCBI GI
                  g4337187
BLAST score
                  297
E value
                  8.0e-27
Match length
                  80
% identity
                   65
                   (AC006403) putative prolylcarboxypeptidase, 5' partial
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   437
                   247 2.R1040
Contig ID
5'-most EST
                  LIB3170-041-Q1-K1-F1
Seq. No.
                   438
                   248 1.R1040
Contig ID
5'-most EST
                  q5509667
Method
                  BLASTN
NCBI GI
                   g403328
BLAST score
                   59
E value
                   3.0e-24
Match length
                   95
% identity
                   91
                  T.repens TrMT1B mRNA for metallothionein-like protein
NCBI Description
Seq. No.
                   439
                   248 2.R1040
Contig ID
5'-most EST
                   q4300393
                  BLASTN
Method
NCBI GI
                   g403328
BLAST score
                   56
                   1.0e-22
E value
Match length
                   115
% identity
                   91
NCBI Description
                  T.repens TrMT1B mRNA for metallothionein-like protein
Seq. No.
                   440
                   248 3.R1040
Contig ID
5'-most EST
                   g4283686
Method
                   BLASTN
NCBI GI
                   g403328
BLAST score
                   56
E value
                   1.0e-22
                   92
Match length
                   90
% identity
NCBI Description T.repens TrMT1B mRNA for metallothionein-like protein
                   441
Seq. No.
Contig ID
                   249 1.R1040
5'-most EST
                   LIB3073-012-Q1-K1-H3
Method
                   BLASTX
```

g4574320

311 4.0e-28



```
107
Match length
% identity
NCBI Description
                  (AF117224) wound-induced protein WI12 [Mesembryanthemum
                  crystallinum]
                  442
Seq. No.
                  249 2.R1040
Contig ID
5'-most EST
                  zzp700831257.h1
Method
                  BLASTX
NCBI GI
                  g4574320
BLAST score
                  306
E value
                  2.0e-27
Match length
                  107
                  60
% identity
                  (AF117224) wound-induced protein WI12 [Mesembryanthemum
NCBI Description
                  crystallinum]
                  443
Seq. No.
                  249 3.R1040
Contig ID
5'-most EST
                  rca700999690.h1
Seq. No.
                  444
Contig ID
                  250 1.R1040
5'-most EST
                  jC-gmle01810014c12d1
Method
                  BLASTX
NCBI GI
                  q4468218
BLAST score
                  378
                  6.0e-36
E value
Match length
                  174
% identity
                  44
                  (AJ010025) unr-interacting protein [Homo sapiens]
NCBI Description
                  445
Seq. No.
Contig ID
                  251_1.R1040
5'-most EST
                  fC-gmse700840436_FL
Seq. No.
                  446
                  251 2.R1040
Contig ID
5'-most EST
                  LIB3050-012-Q1-E1-H5
                  447
Seq. No.
                  252 1.R1040
Contig ID
5'-most EST
                  fC-gmse700840647_FL
Method
                  BLASTX
NCBI GI
                  g3292817
BLAST score
                  838
E value
                  1.0e-89
Match length
                  209
% identity
                  (AL031018) hypothetical protein [Arabidopsis thaliana]
NCBI Description
```

Seq. No. 448

252 2.R1040 Contig ID 5'-most EST g4260232 Method BLASTX NCBI GI g3292817 BLAST score 455



```
E value
                   4.0e-45
Match length
                   155
                   59
% identity
NCBI Description
                  (AL031018) hypothetical protein [Arabidopsis thaliana]
                   449
Seq. No.
                   252 3.R1040
Contig ID
5'-most EST
                   LIB3138-011-Q1-N1-B5
Method
                   BLASTX
NCBI GI
                   q3292817
BLAST score
                   276
E value
                   2.0e-24
Match length
                   89
% identity
                   69
                  (AL031018) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   450
                   253 1.R1040
Contig ID
5'-most EST
                   fC-gmse700841225_FL
                   {\tt BLASTX}
Method
NCBI GI
                   g2642430
BLAST score
                   432
E value
                   2.0e-42
                   157
Match length
                   62
% identity
NCBI Description
                  (AC002391) putative AP2 domain containing protein
                   [Arabidopsis thaliana]
Seq. No.
                   451
Contig ID
                   254 1.R1040
5'-most EST
                   LIB3049-015-Q1-E1-C2
Method
                   BLASTX
                   q2894605
NCBI GI
BLAST score
                   267
E value
                   5.0e-23
Match length
                   133
% identity
                   50
                  (AL021889) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   452
                   254 2.R1040
Contig ID
5'-most EST
                   hrw701059629.h1
Method
                   BLASTX
NCBI GI
                   g2894605
BLAST score
                   301
E value
                   3.0e-27
Match length
                   131
                   55
% identity
                  (AL021889) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   453
```

Contig ID

255 1.R1040 5'-most EST dpv701100539.h1

Method BLASTX NCBI GI q3335372 BLAST score 1134 E value 1.0e-124



368 Match length % identity 59

NCBI Description (AC003028) putative SRG1 protein [Arabidopsis thaliana]

Seq. No. 454

255 2.R1040 Contig ID

5'-most EST fC-gmro700869137f3

Method BLASTX g3335372 NCBI GI BLAST score 682 E value 2.0e-71 Match length 305 % identity 59

(AC003028) putative SRG1 protein [Arabidopsis thaliana] NCBI Description

Seq. No. 455

Contig ID 255 3.R1040 5'-most EST LIB3040-005-Q1-E1-G2

Method BLASTX NCBI GI q3335372 BLAST score 401 E value 7.0e-39

Match length 165 % identity 50

(AC003028) putative SRG1 protein [Arabidopsis thaliana] NCBI Description

456 Seq. No.

Contig ID 255 4.R1040

5'-most EST LIB3049-047-Q1-E1-F9

Method BLASTX NCBI GI q3335372 BLAST score 150 E value 7.0e-20 Match length 141 % identity 41

NCBI Description (AC003028) putative SRG1 protein [Arabidopsis thaliana]

457 Seq. No.

255 5.R1040 Contig ID

5'-most EST LIB3040-016-Q1-E1-H5

BLASTX Method NCBI GI g3335372 BLAST score 113 E value 1.0e-12 Match length 84 48 % identity

NCBI Description (AC003028) putative SRG1 protein [Arabidopsis thaliana]

458 Seq. No.

Contig ID 256 1.R1040 5'-most EST asn701139845.h1

Method BLASTX NCBI GI g2702268 BLAST score 655 6.0e-77 E value Match length 254 % identity 61



NCBI Description (AC003033) putative cellulase [Arabidopsis thaliana]

Seq. No. 459

257 1.R1040 Contig ID

jC-gmro02910064g02a1 5'-most EST

BLASTX Method g1518540 NCBI GI BLAST score 2252 0.0e + 00E value Match length 480 89 % identity

NCBI Description (U53418) UDP-glucose dehydrogenase [Glycine max]

460 Seq. No.

257 2.R1040 Contig ID

fC-gmf1700899609f1 5'-most EST

Method BLASTX g3378650 NCBI GI BLAST score 862 E value 9.0e-93 239 Match length 70 % identity

NCBI Description (X97606) abscisic acid activated [Medicago sativa]

Seq. No. 461

Contig ID 257 3.R1040

5'-most EST LIB3092-040-Q1-K1-H3

Seq. No. 462

257 4.R1040 Contig ID 5'-most EST ssr700556162.hl

463 Seq. No.

Contig ID 257 5.R1040 5'-most EST g4314063

Seq. No. 464

257 6.R1040 Contig ID

5'-most EST jC-gmf102220146e08d1

BLASTX Method NCBI GI g3378650 BLAST score 215 E value 3.0e-17 63 Match length % identity

NCBI Description (X97606) abscisic acid activated [Medicago sativa]

465 Seq. No.

257 9.R1040 Contig ID

jC-gmf102220114g07a15'-most EST

466 Seq. No.

257 10.R1040 Contig ID

5'-most EST uC-gmflminsoy047d09b1

Seq. No.

Contig ID 257 11.R1040

Match length

247



```
5'-most EST
                  jC-qmst02400056g02a1
Seq. No.
                  468
                  258 1.R1040
Contig ID
5'-most EST
                  jC-gmst02400062d08d1
Method
                  BLASTX
NCBI GI
                  g2146739
BLAST score
                  1772
E value
                  0.0e + 00
Match length
                  469
                  72
% identity
                  hexokinase (EC 2.7.1.1) 1 - Arabidopsis thaliana >gi_881521
NCBI Description
                  (U28214) hexokinase 1 [Arabidopsis thaliana]
Seq. No.
                  469
Contig ID
                  258 2.R1040
5'-most EST
                  jC-gmro02910073c05d1
Method
                  BLASTX
NCBI GI
                  q1899025
BLAST score
                  692
E value
                  1.0e-72
Match length
                  185
% identity
                  71
                  (U28215) hexokinase 2 [Arabidopsis thaliana] >gi 3687232
NCBI Description
                  (AC005169) hexokinase [Arabidopsis thaliana]
Seq. No.
                  470
Contig ID
                  258 3.R1040
5'-most EST
                  jC-gmf102220054c06a1
Method
                  BLASTX
                  g2146739
NCBI GI
BLAST score
                  267
E value
                  3.0e-23
Match length
                  68
% identity
                  78
                  hexokinase (EC 2.7.1.1) 1 - Arabidopsis thaliana >gi 881521
NCBI Description
                  (U28214) hexokinase 1 [Arabidopsis thaliana]
                  471
Seq. No.
                  259 1.R1040
Contig ID
5'-most EST
                  jC-gmro02910035d12d1
Method
                  BLASTX
NCBI GI
                  g3702332
BLAST score
                  342
E value
                  1.0e-31
Match length
                  193
% identity
NCBI Description
                  (AC005397) unknown protein [Arabidopsis thaliana]
Seq. No.
                  472
                  260 1.R1040
Contig ID
5'-most EST
                  fC-zmro700830504 FL
Method
                  BLASTX
NCBI GI
                  q4240116
BLAST score
                  1159
E value
                  1.0e-127
```



% identity

NCBI Description

(AB007799) NADH-cytochrome b5 reductase [Arabidopsis thaliana] >gi_4240118_dbj_BAA74838_ (AB007800)

NADH-cytochrome b5 reductase [Arabidopsis thaliana]

Seq. No.

Contig ID 5'-most EST 260 2.R1040 g4396819

Seq. No.

Contig ID

261 1.R1040

5'-most EST

fC-zmro700830907_FL

Method NCBI GI BLASTX g4191788 768

E value Match length

BLAST score

1.0e-81 191 76

475

474

% identity NCBI Description

(AC005917) putative 1-aminocyclopropane-1-carboxylate

oxidase [Arabidopsis thaliana]

Seq. No.

Contig ID

262 1.R1040

5'-most EST

LIB3051-090-Q1-K1-G3

Method NCBI GI BLASTN q1336096 581

BLAST score E value Match length

0.0e+001160

% identity NCBI Description

88 Pisum sativum pyruvate dehydrogenase Elbeta mRNA, nuclear

gene encoding mitochondrial protein, complete cds

Seq. No.

Contig ID

262 2.R1040

5'-most EST

uC-gmropic021f10b1

Method NCBI GI BLASTN g1336096 153

BLAST score E value Match length

2.0e-80 347

% identity

NCBI Description

Pisum sativum pyruvate dehydrogenase Elbeta mRNA, nuclear

gene encoding mitochondrial protein, complete cds

Seq. No.

477

Contig ID

262 3.R1040

5'-most EST

LIB3051-081-Q1-K1-E5

Method NCBI GI BLASTN g1336096

BLAST score

E value

133 2.0e-68

Match length % identity

335 85

NCBI Description

Pisum sativum pyruvate dehydrogenase Elbeta mRNA, nuclear

gene encoding mitochondrial protein, complete cds



Method BLASTX
NCBI GI g1709454
BLAST score 230
E value 2.0e-19
Match length 63

% identity 78
NCBI Description PYRUVATE DEHYDROGENASE E1 COMPONENT, BETA SUBUNIT PRECURSOR

(PDHE1-B) >gi_1336097 (U56697) pyruvate dehydrogenase

Elbeta [Pisum sativum]

Seq. No. 479

Contig ID 264_1.R1040

5'-most EST fC-zmro700832582_FL

Method BLASTX
NCBI GI g2959732
BLAST score 353
E value 3.0e-33
Match length 71
% identity 89

NCBI Description (Y13649) homologous to GATA-binding transcription factors

[Arabidopsis thaliana]

Seq. No. 480

Contig ID 265_1.R1040

5'-most EST fC-zmro700833374_FL

Method BLASTX
NCBI GI g1872521
BLAST score 303
E value 4.0e-27
Match length 113
% identity 36

NCBI Description (U87833) zinc-finger protein Lsd1 [Arabidopsis thaliana]

>gi_1872523 (U87834) zinc-finger protein Lsd1 [Arabidopsis

thaliana]

Seq. No. 481

Contig ID 267 1.R1040

5'-most EST LIB3052-002-Q1-B1-G3

Method BLASTX
NCBI GI g1942405
BLAST score 4443
E value 0.0e+00
Match length 857
% identity 98

NCBI Description Lipoxygenase-3(Soybean) Non-Heme Fe(Ii) Metalloprotein

>gi_4388888_pdb_1BYT_ Lipoxygenase-3(Soybean) Complex With

4-Nitrocatechol >gi_1794172 (U50081) lipoxygenase-3

[Glycine max]

Seq. No. 482

Contig ID 268_2.R1040

5'-most EST fC-gmro700568375r1

Method BLASTX NCBI GI g1931639



```
BLAST score
                    3.0e - 31
 E value
Match length
                    148
                    53
 % identity
 NCBI Description (U95973) lysophospholipase isolog [Arabidopsis thaliana]
                    483
 Seq. No.
                    268 3.R1040
 Contig ID
                    LIB3074-031-Q1-K2-E10
 5'-most EST
                    BLASTX
 Method
                    q2832681
 NCBI GI
                    199
 BLAST score
                    7.0e-29
 E value
                    71
 Match length
 % identity
                    (AL021712) putative protein [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
 Contig ID
                    268 4.R1040
                    awf700841481.h1
 5'-most EST
                    BLASTN
 Method
                    g1885372
 NCBI GI
 BLAST score
                    601
 E value
                    0.0e + 00
                    641
 Match length
 % identity
                    99
                    Glycine max small subunit ribosomal RNA gene, partial
 NCBI Description
                    sequence, internal transcribed spacer 1, 5.8S ribosomal RNA
                    gene and internal transcribed spacer 2, complete sequence,
                    and large subunit ribosomal RNA gene, partial seq
                    485
  Seq. No.
                    268 5.R1040
  Contig ID
                    asn701138178.h1
  5'-most EST
                    BLASTX
  Method
                    g3157928
  NCBI GI
                    911
  BLAST score
                    1.0e-155
 E value
                    404
  Match length
  % identity
                    (AC002131) Similar to fumarylacetoacetate hydrolase,
  NCBI Description
                    gb_L41670 from Emericella nidulans. [Arabidopsis thaliana]
                     486
  Seq. No.
                    268 6.R1040
  Contig ID
                    LIB3107-080-Q1-K1-D2
  5'-most EST
                    BLASTN
  Method
                     q170019
  NCBI GI
  BLAST score
                     599
                     0.0e+00
  E value
                     1094
  Match length
                     95
  % identity
                    Soybean maturation protein (MAT1) gene, complete cds
  NCBI Description
                     487
  Seq. No.
```

268 7.R1040

uxk700672229.h1

Contig ID

5'-most EST



Method BLASTN
NCBI GI 9497416
BLAST score 628
E value 0.0e+00
Match length 912
% identity 94

NCBI Description Glycine max Essex dehydrin-like protein mRNA, complete cds

Seq. No. 488

Contig ID 269_1.R1040 5'-most EST leu701154837.h1

Method BLASTN
NCBI GI g20831
BLAST score 275
E value 1.0e-153
Match length 791
% identity 84

NCBI Description P.sativum petC mRNA for chloroplast Rieske FeS protein

Seq. No. 489

Contig ID 270 1.R1040

5'-most EST fC-gmst7007923230 FL

Method BLASTX
NCBI GI g4490737
BLAST score 765
E value 4.0e-81
Match length 284
% identity 59

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 490

Contig ID 270_2.R1040

5'-most EST LIB3039-004-Q1-E1-F9

Method BLASTX
NCBI GI g4490737
BLAST score 266
E value 3.0e-23
Match length 87
% identity 62

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 491

Contig ID 270_3.R1040 5'-most EST dpv701097092.h1

Method BLASTX
NCBI GI 94490737
BLAST score 155
E value 2.0e-10
Match length 55
% identity 60

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 492

Contig ID 271_1.R1040

5'-most EST fC-zmf1700381665_FL

Method BLASTX NCBI GI g2129550



BLAST score 1218 1.0e-134 E value Match length 294 76 % identity

NCBI Description calcium-dependent protein kinase (EC 2.7.1.-) CDPK6 -

Arabidopsis thaliana >gi_2129554_pir__S71901 calcium-dependent protein kinase 6 - Arabidopsis thaliana

>gi_836940 (U20623) calcium-dependent protein kinase

[Arabidopsis thaliana] >gi_836944 (U20625)

calcium-dependent protein kinase [Arabidopsis thaliana] >gi_4454034_emb_CAA23031.1_ (AL035394) calcium-dependent

protein kinase (CDPK6) [Arabidopsis thaliana]

Seq. No. 493

272 1.R1040 Contig ID

5'-most EST fC-gmse700764850 FL

Method BLASTX NCBI GI g4417267 BLAST score 315 E value 1.0e-28 Match length 152 46 % identity

NCBI Description (AC007019) hypothetical protein [Arabidopsis thaliana]

Seq. No.

494

Contig ID 273 3.R1040

5'-most EST LIB3051-003-Q1-E1-D8

Seq. No. 495

273 4.R1040 Contig ID 5'-most EST g5676977

Seq. No.

Contig ID 273 5.R1040

5'-most EST jC-gmf102220132ag03d1

496

Seq. No. 497

Contig ID 273 6.R1040

5'-most EST LIB3051-067-Q1-K1-A9

Method BLASTX NCBI GI g2911280 BLAST score 482 E value 2.0e-48 Match length 107 % identity 87

NCBI Description (U73937) PK12 protein kinase [Nicotiana tabacum]

Seq. No. 498

Contig ID 273 7.R1040 5'-most EST seb700650989.hl

Method BLASTX NCBI GI g133438 BLAST score 677 E value 4.0e-71 Match length 180 % identity 74

NCBI Description DNA-DIRECTED RNA POLYMERASE BETA' CHAIN



>gi_81504_pir__B29959 DNA-directed RNA polymerase (EC
2.7.7.6) beta' chain - spinach chloroplast >gi_295120
(M55297) RNA polymerase [Spinacia oleracea]

Seq. No. 499
Contig ID 273_8.R1040
5'-most EST LIB3040-060-Q1-E1-H9
Method BLASTX
NCBI GI g3548802
BLAST score 553
E value 2.0e-56

E value 2.0e Match length 243 % identity 44

NCBI Description (AC005313) axi 1-like protein [Arabidopsis thaliana]

>gi_4335769_gb_AAD17446_ (AC006284) putative axi1 protein

[Nicotiana tabacum] [Arabidopsis thaliana]

Seq. No. 500

Contig ID 273 9.R1040

5'-most EST LIB3109-052-Q1-K1-F3

Method BLASTX
NCBI GI g2894568
BLAST score 173
E value 4.0e-12
Match length 93
% identity 46

NCBI Description (AL021890) putative protein [Arabidopsis thaliana]

Seq. No. 501

Contig ID 273 10.R1040

5'-most EST LIB3049-042-Q1-E1-E8

Seq. No. 502

Contig ID 273 11.R1040

5'-most EST LIB3049-041-Q1-E1-H3

Seq. No. 503

Contig ID 273_12.R1040

5'-most EST LIB3092-031-Q1-K1-H9

Seq. No. 504

Contig ID 273_13.R1040 5'-most EST hrw701060522.h1

Method BLASTX
NCBI GI g294845
BLAST score 744
E value 1.0e-78
Match length 261
% identity 56

NCBI Description (L13655) membrane protein [Saccharum hybrid cultivar

H65-7052]

Seq. No. 505

Contig ID 273 14.R1040

5'-most EST jC-gmle01810014h09d1

Method BLASTX NCBI GI g2911280



BLAST score E value 2.0e-67 Match length 197 % identity 69

(U73937) PK12 protein kinase [Nicotiana tabacum] NCBI Description

Seq. No. Contig ID

506 273 15.R1040 5'-most EST jex700906210.h1

Seq. No.

507

273 16.R1040 Contig ID

5'-most EST LIB3051-106-Q1-K1-E6

Method BLASTN NCBI GI g2760164 BLAST score 52 E value 1.0e-19 Match length 182 % identity 86

Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: NCBI Description

K18P6, complete sequence [Arabidopsis thaliana]

508 Seq. No.

Contig ID 273 17.R1040

5'-most EST LIB3170-049-Q1-K1-D12

Method BLASTX NCBI GI g3643249 BLAST score 545 E value 1.0e-55 Match length 124 75 % identity

(AF090143) thaumatin-like protein precursor Mdtl1 [Malus NCBI Description

domestica]

Seq. No. 509

Contig ID 273 18.R1040

5'-most EST LIB3170-037-Q1-J1-A8

Method BLASTN NCBI GI q22073 BLAST score 124 E value 1.0e-62 Match length 144 % identity 97

NCBI Description Mung bean 25S rRNA - 18S rDNA spacer region

Seq. No. 510

273 19.R1040 Contig ID

5'-most EST LIB3028-001-Q1-B1-G12

Method BLASTN NCBI GI q1675195 BLAST score 99 E value 6.0e-48 Match length 316 86 % identity

NCBI Description Pisum sativum (clone PsRCI22-3) callus protein P23 mRNA,

complete cds



```
Seq. No.
                   511
                  273 20.R1040
Contig ID
                  jsh701065521.h1
5'-most EST
```

Method BLASTX q1916809 NCBI GI BLAST score 483 1.0e-48 E value 148 Match length % identity 64

NCBI Description (U81163) auxin-binding protein [Prunus persica]

Seq. No. 512

273 21.R1040 Contig ID

5'-most EST LIB3109-028-Q1-K1-C6

Method BLASTX NCBI GI g1703200 BLAST score 948 E value 1.0e-102 212 Match length % identity 82

PROTEIN KINASE AFC2 >gi 601789 (U16177) protein kinase NCBI Description [Arabidopsis thaliana] >gi 642130 dbj BAA08214 (D45353)

protein kinase [Arabidopsis thaliana]

>gi_4220516_emb_CAA22989_ (AL035356) protein kinase (AFC2)

[Arabidopsis thaliana]

Seq. No. 513

Contig ID 273 22.R1040

5'-most EST LIB3109-002-Q1-K3-G12

Method BLASTX NCBI GI g4337027 BLAST score 1473 E value 1.0e-178 Match length 687

% identity 72

NCBI Description (AF123254) MFP2 [Arabidopsis thaliana]

514 Seq. No.

Contig ID 273 23.R1040

5'-most EST LIB3138-064-Q1-N1-B5

Method BLASTN NCBI GI g2656024 BLAST score 50 2.0e-18 E value Match length 74 % identity 92

Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: NCBI Description

K15E6

515 Seq. No.

Contig ID 273 24.R1040 5'-most EST leu701145525.h1

Method BLASTX NCBI GI g4193388 BLAST score 182 6.0e-20 E value Match length 171



% identity 29
NCBI Description (AF091455) translationally controlled tumor protein [Hevea brasiliensis]

Seq. No. 516 273 25.R1040 Contig ID q4313694 5'-most EST Method BLASTX NCBI GI q4006890 BLAST score 472 E value 2.0e-46 Match length 112 % identity 78

NCBI Description (Z99708) ubiquitin--protein ligase-like protein

[Arabidopsis thaliana]

Seq. No. 517

Contig ID 273 26.R1040

5'-most EST LIB3139-062-P1-N1-F8

Method BLASTN
NCBI GI g210811
BLAST score 490
E value 0.0e+00
Match length 1716
% identity 86

NCBI Description Bean pod mottle virus coat protein gene, complete cds,

complete middle component (M) RNA

Seq. No. 518

Contig ID 274 1.R1040

5'-most EST fC-gmse700606076 FL

Method BLASTX
NCBI GI g1688233
BLAST score 415
E value 2.0e-40
Match length 191
% identity 57

NCBI Description (U77655) DNA binding protein homolog [Solanum tuberosum]

Seq. No. 519

Contig ID 275 1.R1040

5'-most EST fC-gmst700648235_FL

Method BLASTX
NCBI GI g544134
BLAST score 722
E value 5.0e-76
Match length 353
% identity 47

NCBI Description DNA-DAMAGE-REPAIR/TOLERATION PROTEIN DRT100 PRECURSOR

>gi 99720_pir S22863 hypothetical protein - Arabidopsis
thaliana >gi 421844 pir A46260 RecA functional analog

DRT100 - Arabidopsis thaliana (fragment)

Seq. No. 520

Contig ID 276 1.R1040

5'-most EST fC-qmst700658616 FL

Method BLASTX



```
q2497540
NCBI GI
BLAST score
                  646
E value
                  2.0e-67
Match length
                  183
% identity
                  69
NCBI Description PYRUVATE KINASE, CHLOROPLAST ISOZYME G
Seq. No.
                  521
Contig ID
                  277 1.R1040
5'-most EST
                  LIB3167-027-P1-K1-F8
Method
                  BLASTX
NCBI GI
                  g2144584
BLAST score
                  522
E value
                  7.0e-53
Match length
                  198
                  56
% identity
                 trypsin inhibitor A (Kunitz) precursor - soybean
NCBI Description
                  >gi 18770 emb CAA45777 (X64447) trypsin inhibitor subtype
                  A [Glycine max]
Seq. No.
                  522
Contig ID
                  278 1.R1040
5'-most EST
                  fC-qmst700650610i2
Method
                  BLASTX
NCBI GI
                  g1362078
BLAST score
                  1082
E value
                  1.0e-118
Match length
                  286
% identity
                  67
NCBI Description
                  endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG1)
                  - common nasturtium >gi 311835 emb CAA48324 (X68254)
                  cellulase [Tropaeolum majus]
Seq. No.
                  523
                  278 2.R1040
Contig ID
                  LIB3106-099-Q1-K1-C4
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1084391
BLAST score
                  472
                  6.0e-47
E value
                  124
Match length
                  68.
% identity
                  endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG2)
NCBI Description
                  - common nasturtium (fragment) >gi_311837_emb_CAA48325_
                   (X68255) cellulase [Tropaeolum majus]
Seq. No.
                  524
Contig ID
                  278 3.R1040
5'-most EST
                  seb700650991.h1
Method
                  BLASTX
NCBI GI
                  g1362078
```

Method BLASTX
NCBI GI g1362078
BLAST score 197
E value 2.0e-15
Match length 49
% identity 69

NCBI Description endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG1) - common nasturtium >gi 311835 emb CAA48324 (X68254)



cellulase [Tropaeolum majus]

525 Seq. No. 279 1.R1040 Contig ID

fC-gmse700645511 FL 5'-most EST

BLASTX Method NCBI GI q3820648 BLAST score 1471 E value 0.0e + 00Match length 469 % identity 67

(Y12636) allene oxide synthase [Arabidopsis thaliana] NCBI Description

Seq. No.

279_2.R1040 Contig ID

5'-most EST jC-gmro02910032b07d1

Method BLASTX NCBI GI g1890152 BLAST score 520 7.0e-53 E value 122 Match length % identity 77

(X92510) allene oxide synthase [Arabidopsis thaliana] NCBI Description

Seq. No. 527

Contig ID 279 3.R1040

5'-most EST jC-gmro02910032e06a1

Method BLASTX NCBI GI q1352186 BLAST score 406 1.0e-39 E value Match length 141 % identity 61

ALLENE OXIDE SYNTHASE PRECURSOR (HYDROPEROXIDE DEHYDRASE) NCBI Description

(CYTOCHROME P450 74) >gi 404866 (U00428) allene oxide

synthase [Linum usitatissimum]

528 Seq. No.

279 4.R1040 Contig ID

5'-most EST LIB3139-066-P1-N1-A9

Method BLASTX NCBI GI g1352186 BLAST score 234 E value 2.0e-19 89 Match length 56 % identity

ALLENE OXIDE SYNTHASE PRECURSOR (HYDROPEROXIDE DEHYDRASE) NCBI Description

(CYTOCHROME P450 74) $>gi_404866$ (U00428) allene oxide

synthase [Linum usitatissimum]

529 Seq. No.

Contig ID 280 1.R1040

5'-most EST fC-gmse700657716 FL

Seq. No. 530

Contig ID 281 1.R1040

5'-most EST fC-gmse700657751 FL



Method BLASTX
NCBI GI g4006895
BLAST score 423
E value 2.0e-41
Match length 139
% identity 55

NCBI Description (Z99708) putative protein [Arabidopsis thaliana]

Seq. No. Contig ID

531 282 1.R1040

5'-most EST

fC-gmse700657823 FL

Method BLASTX
NCBI GI g3157923
BLAST score 208
E value 1.0e-16
Match length 71

Match length 71 % identity 63

NCBI Description (AC002131) F12F1.7 [Arabidopsis thaliana]

Seq. No. Contig ID

NO. 53.

283 1.R1040

5'-most EST

fC-gmse700657837_FL

Seq. No.

533

Contig ID

284 1.R1040

5'-most EST

fC-gmse700658078_FL

Seq. No.

534

Contig ID

285 1.R1040

5'-most EST

fC-gmse700658166_FL

Seq. No.

535

Contig ID

286_1.R1040 pst700645867.h1

5'-most EST Method

BLASTX

NCBI GI BLAST score

g4510402 806

E value Match length 6.0e-86

% identity

311 58

NCBI Description

(AC006587) putative AP2 domain [Arabidopsis thaliana]

Seq. No.

536

Contig ID

287 1.R1040

5'-most EST

fC-gmse700660136_FL BLASTX

Method NCBI GI

g1531760

BLAST score

1055

E value Match length 1.0e-115

% identity

329

NCBI Description

(X97075) proline oxidase [Arabidopsis thaliana]

Seq. No.

537

Contig ID

287 2.R1040

5'-most EST

LIB3170-037-Q1-J1-A5

Method

BLASTX



```
q1531760
NCBI GI
                  295
BLAST score
                  1.0e-26
E value
Match length
                  86
                  69
% identity
                  (X97075) proline oxidase [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                  288 1.R1040
                  LIB3073-014-Q1-K1-F8
5'-most EST
Seq. No.
                   539
                   289 1.R1040
Contig ID
                   fC-gmse700661496 FL
5'-most EST
Method
                   BLASTX
                   g2983642
NCBI GI
                   732
BLAST score
                   4.0e-77
E value
                   409
Match length
% identity
NCBI Description (AE000728) diaminopimelate decarboxylase [Aquifex aeolicus]
                   540
Seq. No.
                   290 1.R1040
Contig ID
                   uC-gmrominsoy258c06b1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4335756
BLAST score
                   833
                   3.0e-89
E value
Match length
                   229
% identity
                   72
                  (AC006284) putative ankyrin [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   291 1.R1040
Contig ID
                   fC-gmst700651107h1
5'-most EST
                   BLASTX
Method
                   g2499611
NCBI GI
BLAST score
                   503
                   3.0e-98
E value
                   213
Match length
                   83
% identity
                   MITOGEN-ACTIVATED PROTEIN KINASE HOMOLOG 7 (MAP KINASE 7)
NCBI Description
                   (ATMPK7) >gi_629548_pir__S40473 mitogen-activated protein
                   kinase 7 (EC 2.7.1.-) - Arabidopsis thaliana
                   >gi_457406_dbj_BAA04870_ (D21843) MAP kinase [Arabidopsis
                   thaliana]
Seq. No.
                   542
                   292 1.R1040
Contig ID
                   jC-gmro02800039g09d1
5'-most EST
                   BLASTX
Method
                   q399046
NCBI GI
BLAST score
                   557
                   6.0e-57
E value
Match length
                   141
```

82

% identity





```
NCBI Description ADENINE PHOSPHORIBOSYLTRANSFERASE 1 (APRT)
                  >gi_99657_pir__S20867 adenine phosphoribosyltransferase (EC
                  2.4.2.7) - Arabidopsis thaliana >gi 16164 emb CAA41497
                  (X58640) adenine phosphoribosyltransferase [Arabidopsis
                  thaliana] >gi_433050 (L19637) adenine
                  phosphoribosyltransferase [Arabidopsis thaliana]
                  >gi 3935182 (AC004557) F17L21.25 [Arabidopsis thaliana]
Seq. No.
                  543
                  292 2.R1040
Contig ID
                  q4437087
5'-most EST
                  BLASTX
Method
                  g399046
NCBI GI
BLAST score
                  714
                  2.0e-75
E value
Match length
                  160
                  86
% identity
                  ADENINE PHOSPHORIBOSYLTRANSFERASE 1 (APRT)
NCBI Description
                  >gi_99657_pir__S20867 adenine phosphoribosyltransferase (EC
                  2.4.2.7) - Arabidopsis thaliana >gi_16164_emb_CAA41497
                   (X58640) adenine phosphoribosyltransferase [Arabidopsis
                   thaliana] >gi 433050 (L19637) adenine
                   phosphoribosy\overline{1}transferase [Arabidopsis thaliana]
                  >gi_3935182 (AC004557) F17L21.25 [Arabidopsis thaliana]
Seq. No.
                   292 3.R1040
Contig ID
                   g4313406
5'-most EST
Seq. No.
                   545
                   293 1.R1040
Contig ID
                   LIB3028-030-Q1-B1-A3
5'-most EST
                   BLASTX
Method
                   g2143629
NCBI GI
                   227
BLAST score
                   3.0e-18
E value
                   172
Match length
                   35
% identity
                   Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV
NCBI Description
                   - rat >gi 986941 (L42810) 'Ca2+/calmodulin-dependent
                   protein kinase kinase' [Rattus norvegicus]
                   >gi_1583255_prf__2120334A Ca/calmodulin-dependent protein
                   kinase kinase [Rattus norvegicus]
                   546
Seq. No.
                   293 2.R1040
Contig ID
                   zhf700964957.h1
 5'-most EST
                   BLASTX
Method
NCBI GI
                   q2143629
```

BLAST score 445

1.0e-43 E value 254 Match length 42 % identity NCBI Description

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - rat >gi_986941 (L42810) 'Ca2+/calmodulin-dependent

24

protein kinase kinase' [Rattus norvegicus]

>gi_1583255_prf__2120334A Ca/calmodulin-dependent protein

Seq. No.

551



kinase kinase [Rattus norvegicus]

```
Seq. No.
                   547
                   293 3.R1040
Contig ID
5'-most EST
                   uC-gmronoir074f07b1
Method
                   BLASTX
NCBI GI
                   g3024670
BLAST score
                   206
E value
                   7.0e-16
Match length
                   126
                   35
% identity
NCBI Description
                   SERINE/THREONINE-PROTEIN KINASE 11
                   (SERINE/THREONINE-PROTEIN KINASE LKB1) >gi 1480861 (U63333)
                   serine/threonine protein kinase [Homo sapiens] >gi 2754827
                   (AF035625) serine threonine kinase 11 [Homo sapiens]
                   >gi 2791676 (AF032986) serine threonine kinase 11 [Homo
                   sapiens] >gi_3063585 (AF055327) Peutz-Jeghers syndrome
                   protein [Homo sapiens] >gi_4507271_ref_NP_000446.1_pSTK11_
serine/threonine kinase 11 (Peutz-Jeghers syndrome)
Seq. No.
                   548
                   294 1.R1040
Contig ID
5'-most EST
                   uC-gmropic042b12b1
Method
                   BLASTX
NCBI GI
                   g3790188
BLAST score
                   1521
E value
                   1.0e-170
Match length
                   360
% identity
                   82
NCBI Description
                   (Y14431) NAD-dependent isocitrate dehydrogenase [Nicotiana
                   tabacum]
                   549
Seq. No.
Contig ID
                   294 3.R1040
5'-most EST
                   epx701105402.h1
Method
                   BLASTX
NCBI GI
                   g3790188
BLAST score
                   369
E value
                   2.0e-35
Match length
                   91
                   78
% identity
                   (Y14431) NAD-dependent isocitrate dehydrogenase [Nicotiana
NCBI Description
                   tabacum]
                   550
Seq. No.
                   294 4.R1040
Contig ID
5'-most EST
                   fC-gmse700839576r1
Method
                   BLASTX
NCBI GI
                   g3080424
BLAST score
                   142
E value
                   7.0e-09
Match length
                   41
% identity
NCBI Description
                   (AL022604) NAD+ dependent isocitrate dehydrogenase subunit
                   1 [Arabidopsis thaliana]
```



```
295 1.R1040
Contig ID
                   fC-gmst700661828 FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   g119791
                   1077
BLAST score
                   1.0e-117
E value
                   319
Match length
                   71
% identity
                   3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE PRECURSOR
NCBI Description
                   (3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)
                   >gi_1084385_pir__S22450 3-oxoacyl-[acyl-carrier-protein]
reductase (EC 1.1.1.100) precursor - Cuphea lanceolata
                   >gi 18046 emb_CAA45866 (X64566) 3-oxoacyl-[acyl-carrier
                   protein] reductase [Cuphea lanceolata]
                   >gi 228929_prf__1814446A beta ketoacyl-ACP reductase
                   [Cuphea lanceolata]
                   552
Seq. No.
                   296 1.R1040
Contig ID
                   jsh701066535.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2791947
BLAST score
                   217
                   1.0e-118
E value
Match length
                   616
                   84
% identity
NCBI Description Lupinus luteus mRNA for ribosomal protein L13a
                   553
Seq. No.
                   296 2.R1040
Contig ID
                   hrw701061971.hl
5'-most EST
                   BLASTX
Method
                   q3914535
NCBI GI
                    699
BLAST score
                    8.0e-74
E value
                    168
Match length
                    85
% identity
                    60S RIBOSOMAL PROTEIN L13A >gi 2791948 emb CAA11283
NCBI Description
                    (AJ223363) ribosomal protein L13a [Lupinus luteus]
                    554
Seq. No.
                    296 3.R1040
Contig ID
5'-most EST
                    qsv701053623.hl
Method
                    BLASTN
                    q2791947
NCBI GI
BLAST score
                    186
                    1.0e-100
E value
                    584
Match length
% identity
                    83
                   Lupinus luteus mRNA for ribosomal protein L13a
NCBI Description
Seq. No.
                    555
```

Contig ID

296 4.R1040

LIB3051-088-Q1-K1-E9 5'-most EST

BLASTX Method g3914535 NCBI GI BLAST score 560



E value 1.0e-57 Match length 119 % identity 92

60S RIBOSOMAL PROTEIN L13A >gi 2791948 emb CAA11283 NCBI Description (AJ223363) ribosomal protein L13a [Lupinus luteus]

Seq. No. Contig ID

556 297_1.R1040

5'-most EST

fC-gmst700662785_FL

Method BLASTX NCBI GI q2462749 BLAST score 1006 E value 1.0e-136 Match length 374 % identity 70

(AC002292) Putative Serine/Threonine protein kinase NCBI Description

[Arabidopsis thaliana]

Seq. No.

557

558

Contig ID 298 1.R1040

5'-most EST

fC-gmst700662976 FL

Seq. No.

298 2.R1040 Contig ID 5'-most EST hyd700727184.h1

Seq. No.

559 Contig ID

299 1.R1040

5'-most EST fC-gmst700663110 FL

Method BLASTX NCBI GI q2388581 BLAST score 1062 E value 1.0e-116 Match length 287

% identity

NCBI Description (AC000098) Similar to Zea DWARF3 (gb_U32579). [Arabidopsis

thaliana]

Seq. No.

560

Contig ID 300 1.R1040

5'-most EST fC-gmst700667704 FL

BLASTX Method NCBI GI g169705 BLAST score 1433 E value 1.0e-159 Match length 312

% identity

(M64737) ATP:pyruvate phosphotransferase [Ricinus communis] NCBI Description

Seq. No.

561

Contig ID

301 1.R1040

5'-most EST

LIB3106-093-Q1-K1-D3

Method BLASTX NCBI GI q479386 BLAST score 2286 0.0e+00 E value Match length 441



% identity

NCBI Description isocitrate dehydrogenase - soybean

Seq. No.

562

Contig ID

302 1.R1040

5'-most EST

fC-gmse700658185d3

Method NCBI GI BLASTX g2826842

BLAST score E value

1691 0.0e+00

Match length % identity

491 64

NCBI Description

(AJ002236) loxc homologue [Lycopersicon pimpinellifolium]

Seq. No.

563

Contig ID

302 2.R1040

5'-most EST

fC-gmf1700862926d1

Method NCBI GI BLAST score BLASTX q2826842

E value

366 6.0e-36

Match length % identity

109 70

NCBI Description

(AJ002236) loxc homologue [Lycopersicon pimpinellifolium]

Seq. No.

564

Contig ID

303_1.R1040

5'-most EST

fC_gmse700658351 FL

Method NCBI GI BLASTX g730042

BLAST score E value

154 7.0e-10

Match length % identity

144 31

NCBI Description MOSA PROTEIN >gi 1073154 pir B53308 mosA protein -

Rhizobium meliloti (strain L5-30) >gi 310303 (L17071) mosA

[Rhizobium meliloti]

Seq. No.

565

Contig ID

304 1.R1040

5'-most EST

fC gmf1700348376 FL

Method NCBI GI BLASTX q2760839

BLAST score E value

1044

Match length

1.0e-114

% identity

475

NCBI Description

49

(AC003105) putative receptor kinase [Arabidopsis thaliana]

Seq. No.

566

Contig ID

305_1.R1040

5'-most EST

fC-gmse700658507 FL

Method NCBI GI BLASTN q220729

BLAST score

709

721

E value Match length 0.0e+00



100 % identity

NCBI Description Rat mRNA for cytochrome b5

Seq. No.

567

Contig ID

306 1.R1040

5'-most EST

fC-gmse700658880 FL

Method NCBI GI BLASTX

BLAST score

g3122333

E value

326 6.0e-30

Match length

175

% identity

43

NCBI Description PUTATIVE 3-ISOPROPYLMALATE DEHYDRATASE LARGE SUBUNIT (ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI) >gi 2648330 (AE000954) 3-isopropylmalate dehydratase, large

subunit (leuC) [Archaeoglobus fulgidus]

Seq. No.

568

Contig ID

307 1.R1040

5'-most EST

LIB3106-087-Q1-K1-G7

Method NCBI GI BLASTX

q3894183

BLAST score

432

E value

3.0e-42

Match length

% identity

212 50

NCBI Description (AC005662) calmodulin-like protein [Arabidopsis thaliana]

Seq. No.

569

Contig ID

308_1.R1040

5'-most EST

fC-gmse700659152 FL

Method

BLASTX

NCBI GI

g2459435

BLAST score

746

E value

4.0e-79

Match length % identity

188 73

NCBI Description

(AC002332) putative serine carboxypeptidase [Arabidopsis

thaliana]

Seq. No.

570

Contig ID

309 1.R1040

5'-most EST

hrw701061722.h1

Method

BLASTX

NCBI GI

g1561774

BLAST score

1806

E value

0.0e+00

Match length

460

% identity

NCBI Description

(U67426) malate dehydrogenase [Vitis vinifera]

Seq. No.

571

Contig ID

309 2.R1040

5'-most EST

LIB3170-085-Q1-J1-B1

Method

BLASTX

NCBI GI

g3687228

BLAST score

129



E value 4.0e-16 Match length 86 59 % identity

(AC005169) putative malate dehydrogenase [Arabidopsis NCBI Description

thaliana]

572 Seq. No.

310 1.R1040 Contig ID

5'-most EST fC-gmse700659382_FL

Method BLASTX NCBI GI q1435021 BLAST score 253 E value 1.0e-21 Match length 104 % identity 54

NCBI Description (D26575) DNA-binding protein [Daucus carota]

Seq. No.

Contig ID 310 2.R1040 5'-most EST pmv700890809.h1

Method BLASTX NCBI GI q1435021 BLAST score 161 E value 1.0e-13 81 Match length 58 % identity

(D26575) DNA-binding protein [Daucus carota] NCBI Description

Seq. No. 574

Contig ID 311 1.R1040

5'-most EST fC-qmf1700906722f4

Method BLASTX NCBI GI q3877595 BLAST score 379 2.0e-36 E value Match length 158 50 % identity

NCBI Description (Z79696) predicted using Genefinder; Similarity to Human HEP27 protein (TR:G1079566); cDNA EST yk281b1.5 comes from

this gene; cDNA EST yk281b1.3 comes from this gene [Caenorhabditis elegans] >gi_3879853_emb_CAB04734_

predicted using Genefinder; Similarity to Human HEP27 protein (TR:G1079566); cDNA EST yk281b1.5 comes from this

gene; cDNA EST yk281b1.3 comes from this gene

[Caenorhabditis elegans]

575 Seq. No.

311 2.R1040 Contig ID

5'-most EST fC-gmse700659490 FL

Method BLASTX NCBI GI q4105190 BLAST score 214 E value 3.0e-17 Match length 79 56 % identity

NCBI Description (AF044127) peroxisomal short-chain alcohol dehydrogenase

[Homo sapiens]

Method

NCBI GI

BLAST score

BLASTX q2934696

1750



```
576
Seg. No.
                  311 3.R1040
Contig ID
                  zsg701117538.h2
5'-most EST
                  BLASTX
Method
                  g4105190
NCBI GI
                  165
BLAST score
                  1.0e-11
E value
                  94
Match length
                   48
% identity
                   (AF044127) peroxisomal short-chain alcohol dehydrogenase
NCBI Description
                   [Homo sapiens]
                   577
Seq. No.
                   312 1.R1040
Contig ID
                  LIB3106-052-Q1-K1-C2
5'-most EST
                   BLASTX
Method
                   g3218410
NCBI GI
                   1209
BLAST score
                   1.0e-133
E value
                   412
Match length
                   59
% identity
                   (AL023859) putative prolyl-trna synthetase
NCBI Description
                   [Schizosaccharomyces pombe]
                   578
Seq. No.
                   312 2.R1040
Contig ID
                   ssr700553989.hl
5'-most EST
                   BLASTX
Method
                   g731640
NCBI GI
                   152
BLAST score
                   6.0e-10
E value
                   71
Match length
% identity
                   48
                   PUTATIVE PROLYL-TRNA SYNTHETASE YHRO20W (PROLINE--TRNA
NCBI Description
                   LIGASE) (PRORS) >gi_626755_pir__S46774 multifunctional
                   amino acid--tRNA ligase homolog - yeast (Saccharomyces
                   cerevisiae) >gi_500692 (U10399) Yhr020wp [Saccharomyces
                   cerevisiae]
                   579
Seq. No.
Contig ID
                   313 1.R1040
5'-most EST
                   fC-qmse700659735 FL
Method
                   BLASTX
                   q1001478
NCBI GI
BLAST score
                   236
                   2.0e-19
E value
Match length
                   112
 % identity
                   (D63999) hypothetical protein [Synechocystis sp.]
NCBI Description
                   580
 Seq. No.
 Contig ID
                   314 1.R1040
 5'-most EST
                   g2934695 FL
```



```
E value
                   0.0e + 00
Match length
                   333
% identity
                   96
NCBI Description
```

(AB000097) class III acidic endochitinase [Glycine max]

Seq. No. Contig ID 581 314 2.R1040

5'-most EST Method

uC-gmrominsoy253g01b1

BLASTX NCBI GI g1345971 BLAST score 2308 0.0e+00E value Match length 454 % identity 92

NCBI Description

OMEGA-3 FATTY ACID DESATURASE, CHLOROPLAST PRECURSOR >gi_541947_pir__JQ2339 omega-3 fatty acid desaturase (EC 1.14.99.-) GMD - soybean >gi_408792 (L22965) omega-3 fatty

acid desaturase [Glycine soja]

Seq. No.

Contig ID 5'-most EST

314 3.R1040 trc700562720.h1

582

Method BLASTN NCBI GI g408791 BLAST score 174 E value 1.0e-92 465 Match length % identity 93

NCBI Description

Glycine soja chloroplast omega-3 fatty acid desaturase

(Fadd) mRNA, complete cds

Seq. No.

Contig ID 5'-most EST Method

314 4.R1040 g5688021 BLASTN

583

NCBI GI g408791 BLAST score 457 E value 0.0e + 00Match length 473

99 % identity

Glycine soja chloroplast omega-3 fatty acid desaturase NCBI Description (Fadd) mRNA, complete cds

Seq. No. Contig ID 584 314 5.R1040

5'-most EST

LIB3139-060-P1-N1-E3

Method BLASTN NCBI GI q408791 BLAST score 157 E value 8.0e-83 300 Match length 93 % identity

NCBI Description

Glycine soja chloroplast omega-3 fatty acid desaturase

(Fadd) mRNA, complete cds

Seq. No.

585

Contig ID

314 9.R1040



```
zlv700807515.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q408791
BLAST score
                  94
                  2.0e-45
E value
Match length
                  155
                  100
% identity
NCBI Description
                  Glycine soja chloroplast omega-3 fatty acid desaturase
                  (Fadd) mRNA, complete cds
                  586
Seq. No.
                  315 1.R1040
Contig ID
5'-most EST
                  uC-gmropic039f06b1
Method
                  BLASTN
NCBI GI
                  g2055227
BLAST score
                  258
E value
                  1.0e-143
Match length
                  494
% identity
                  100
NCBI Description Glycine max mRNA for SRC1, complete cds
Seq. No.
                  587
Contig ID
                  315 2.R1040
                  uC-gmflminsoy077e06b1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g2055227
BLAST score
                  248
E value
                  1.0e-137
                  456
Match length
                  90
% identity
NCBI Description Glycine max mRNA for SRC1, complete cds
Seq. No.
                  588
Contig ID
                  316 1.R1040
5'-most EST
                  LIB3093-050-Q1-K1-E10
Method
                  BLASTN
NCBI GI
                  g2055229
BLAST score
                  941
E value
                  0.0e + 00
Match length
                  1100
                  100
% identity
NCBI Description Soybean mRNA for SRC2, complete cds
                  589
Seq. No.
                  316 2.R1040
Contig ID
                  LIB3138-048-Q1-N1-D6
5'-most EST
Method
                  BLASTN
NCBI GI
                  g2055229
BLAST score
                  301
E value
                  1.0e-169
Match length
                  397
```

Seq. No. 590

% identity

Contig ID 316_3.R1040 5'-most EST uaw700661162.h1

94

NCBI Description Soybean mRNA for SRC2, complete cds



```
BLASTN
Method
                   g2055229
NCBI GI
                   205
BLAST score
                   1.0e-111
E value
                   718
Match length
                   93
% identity
NCBI Description Soybean mRNA for SRC2, complete cds
Seq. No.
                   317 3.R1040
Contig ID
                   fC-gmse7000757201d1
5'-most EST
                   BLASTN
Method
                   q2605511
NCBI GI
                   161
BLAST score
                   2.0e-85
E value
                   283
Match length
                   89
                   Glycine max mRNA for beta subunit of beta conglycinin,
% identity
NCBI Description
                    complete cds
                    592
Seq. No.
                    317_4.R1040
Contig ID
                    LIB3170-033-Q1-K1-B12
 5'-most EST
                    BLASTN
Method
                    q2605509
 NCBI GI
 BLAST score
                    250
                    1.0e-138
 E value
                    309
 Match length
                    Glycine max mRNA for alpha subunit of beta conglycinin,
                    95
 % identity
 NCBI Description
                    complete cds
                    593
 Seq. No.
                    317 5.R1040
 Contig ID
                    q50<del>5</del>7596
 5'-most EST
                    BLASTN
 Method
                    g2605511
 NCBI GI
                    1041
 BLAST score
                     0.0e + 00
 E value
 Match length
                    1510
                     95
                    Glycine max mRNA for beta subunit of beta conglycinin,
  % identity
 NCBI Description
                     complete cds
                     594
  Seq. No.
                     317 6.R1040
  Contig ID
                     vwf700678652.hl
  5'-most EST
                     BLASTN
  Method
                     q736001
  NCBI GI
                     1339
  BLAST score
                     0.0e + 00
  E value
                     1782
  Match length
                     96
  % identity
                    G.soja (SH1) Gy5 mRNA for glycinin
  NCBI Description
```

Seq. No. 595 Contig ID 317_7.R1040



```
txt700732304.h1
5'-most EST
                  BLASTN
Method
                  g806555
NCBI GI
                  619
BLAST score
                  0.0e + 00
E value
                  1898
Match length
                  95
% identity
NCBI Description G.soja mRNA for glycinin
                  596
Seq. No.
                   317 8.R1040
Contig ID
                  LIB3072-061-Q1-K1-F11
5'-most EST
                  BLASTN
Method
                   g4191813
NCBI GI
                   1327
BLAST score
                   0.0e + 00
E value
                   1706
Match length
% identity
                   Glycine max mRNA for alpha' subunit of beta-conglycinin,
NCBI Description
                   complete cds
                   597
Seq. No.
                   317 9.R1040
Contig ID
                   g5688089
5'-most EST
                   BLASTN
Method
                   q18535
NCBI GI
                   908
BLAST score
                   0.0e + 00
E value
                   1919
Match length
                   93
% identity
                   Soybean mRNA for the alpha subunit of beta-conglycinin
NCBI Description
                   598
Seq. No.
                   319 1.R1040
Contig ID
5'-most EST
                   qsv701051475.h1
Method
                   BLASTN
                   g1944192
NCBI GI
                   1302
BLAST score
                   0.0e+00
E value
                   1339
Match length
                   99
% identity
                   Glycine max mRNA for nodulin 35, complete cds
NCBI Description
Seq. No.
                   321 1.R1040
Contig ID
                   leu701147347.hl
5'-most EST
                   BLASTX
Method
                   q902938
NCBI GI
BLAST score
                   2584
                   0.0e+00
E value
                   496
Match length
 % identity
                   (D50866) beta-amylase [Glycine max]
NCBI Description
                   >gi_2196550_dbj_BAA20453_ (AB004271) beta-amylase [Glycine
                   max]
```

178

600

Seq. No.



```
321 2.R1040
Contig ID
                  k11701209436.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1483177
BLAST score
                  1690
                  0.0e+00
E value
                  730
Match length
% identity
                  49
                  (D86598) antifreeze-like protein (af70) [Picea abies]
NCBI Description
Seq. No.
                  601
                  321 3.R1040
Contig ID
5'-most EST
                  g5342495
Method
                  BLASTX
NCBI GI
                  g1483177
BLAST score
                  323
E value
                  1.0e-29
Match length
                  179
% identity
                  40
                 (D86598) antifreeze-like protein (af70) [Picea abies]
NCBI Description
                  602
Seq. No.
                  321 4.R1040
Contig ID
5'-most EST
                  g4395964
                  BLASTX
Method
NCBI GI
                  g1483177
BLAST score
                  829
                  9.0e-89
E value
Match length
                  318
% identity
                  55
                 (D86598) antifreeze-like protein (af70) [Picea abies]
NCBI Description
                  603
Seq. No.
                  321 6.R1040
Contig ID
5'-most EST
                  LIB3139-056-P1-N1-E11
                  BLASTX
Method
                  g3176874
NCBI GI
BLAST score
                  221
E value
                  4.0e-18
Match length
                  101
% identity
                  50
                  (AF065639) cucumisin-like serine protease [Arabidopsis
NCBI Description
                  thaliana]
                  604
Seq. No.
                  321 7.R1040
Contig ID
5'-most EST
                  jC-gmst02400066f01a2
Method
                  BLASTX
NCBI GI
                  g1483177
BLAST score
                  208
E value
                  3.0e-20
Match length
                  122
% identity
                  (D86598) antifreeze-like protein (af70) [Picea abies]
NCBI Description
```

605

322 1.R1040

Seq. No. Contig ID



5'-most EST LIB3050-001-Q1-E1-A10 BLASTN Method NCBI GI q762878 BLAST score 461 0.0e+00E value Match length 885 88 % identity Psophocarpus tetragonolobus mRNA for chitinase, complete NCBI Description 606 Seq. No. 322 2.R1040 Contig ID 5'-most EST fC-gmse700671501f1 Method BLASTX NCBI GI q2370253 BLAST score 1290 E value 1.0e-142 Match length 575 % identity 49 (Y13273) putative protein kinase [Lycopersicon esculentum] NCBI Description Seq. No. 607 322 3.R1040 Contig ID 5'-most EST jC-gmle01810025e02d1 Method BLASTN NCBI GI q762878 BLAST score 108 E value 1.0e-53 200 Match length 89 % identity Psophocarpus tetragonolobus mRNA for chitinase, complete NCBI Description 608 Seq. No. 322 5.R1040 Contig ID 5'-most EST fC-gmse700852775a1 Method BLASTX NCBI GI g2370253 BLAST score 418 4.0e-41 E value Match length 140 58 % identity (Y13273) putative protein kinase [Lycopersicon esculentum] NCBI Description Seq. No. 609 322 7.R1040 Contig ID 5'-most EST LIB3051-022-Q1-K1-G7 610 Seq. No. Contig ID 323 1.R1040 5'-most EST zsg701120865.h1

BLASTX Method NCBI GI g2463569 BLAST score 2103 E value 0.0e+00Match length 413 % identity 97





NCBI Description (AB007503) squalene synthase [Glycine max]

Seq. No. 611

Contig ID 323_2.R1040 5'-most EST gsv701053241.h1

Method BLASTN
NCBI GI g2463568
BLAST score 287
E value 1.0e-160
Match length 377
% identity 97

NCBI Description Glycine max mRNA for squalene synthase, complete cds

Seq. No. 612

Contig ID 324_1.R1040 5'-most EST epx701105763.h1

Method BLASTX
NCBI GI g2529229
BLAST score 2535
E value 0.0e+00
Match length 511
% identity 96

NCBI Description (AB007907) 6-phosphogluconate dehydrogenase [Glycine max]

Seq. No. 613

Contig ID 324_3.R1040 5'-most EST crh700851941.h1

Method BLASTN
NCBI GI g2529228
BLAST score 71
E value 1.0e-31
Match length 79
% identity 97

NCBI Description Glycine max mRNA for 6-phosphogluconate dehydrogenase,

complete cds

Seq. No. 614

Contig ID 324 4.R1040

5'-most EST jC-gmf102220061h02d1

Method BLASTN
NCBI GI g2529228
BLAST score 90
E value 6.0e-43
Match length 90
% identity 100

NCBI Description Glycine max mRNA for 6-phosphogluconate dehydrogenase,

complete cds

Seq. No. 615

Contig ID 324_5.R1040 5'-most EST trc700561367.h1

Method BLASTN
NCBI GI g2529228
BLAST score 129
E value 3.0e-66
Match length 280
% identity 86



NCBI Description Glycine max mRNA for 6-phosphogluconate dehydrogenase, complete cds

Seq. No. 616

Contig ID 324_6.R1040 5'-most EST epx701109643.h1

Method BLASTN
NCBI GI g2529228
BLAST score 83
E value 5.0e-39
Match length 139
% identity 90

NCBI Description Glycine max mRNA for 6-phosphogluconate dehydrogenase,

complete cds

Seq. No. 617

Contig ID 324_7.R1040 5'-most EST kl1701214664.h1

Method BLASTN
NCBI GI g2529228
BLAST score 102
E value 3.0e-50
Match length 166
% identity 90

NCBI Description Glycine max mRNA for 6-phosphogluconate dehydrogenase,

complete cds

Seq. No. 618

Contig ID 325_1.R1040 5'-most EST fua701043454.h1

Method BLASTX
NCBI GI g1705587
BLAST score 4790
E value 0.0e+00
Match length 967
% identity 95

NCBI Description PHOSPHOENOLPYRUVATE CARBOXYLASE (PEPCASE)

>gi_474020_dbj_BAA03100_ (D13998) phosphoenolpyruvate

carboxylase [Glycine max]

Seq. No. 619

Contig ID 325_2.R1040
5'-most EST g218266_FL
Method BLASTX
NCBI GI g399182
BLAST score 4874
E value 0.0e+00
Match length 967
% identity 97

% identity 97
NCBI Description PHOSPHOENOLPYRUVATE CARBOXYLASE, HOUSEKEEPING ISOZYME

(PEPCASE) >gi_320128_pir__S28428 phosphoenolpyruvate carboxylase (EC 4.1.1.31) - soybean

>gi_218267_dbj_BAA01560_ (D10717) phosphoenolpyruvate

carboxylase [Glycine max]

Seq. No. 620

Contig ID 325_3.R1040



5'-most EST uC-gmropic021f08b1
Method BLASTX
NCBI GI g1705587
BLAST score 4817
E value 0.0e+00
Match length 967
% identity 96

NCBI Description PHOSPHOENOLPYRUVATE CARBOXYLASE (PEPCASE)

>gi_474020_dbj_BAA03100_ (D13998) phosphoenolpyruvate

carboxylase [Glycine max]

Seq. No. 623

Contig ID 325_7.R1040 5'-most EST sat701006212.h2

Method BLASTN
NCBI GI g467551
BLAST score 212
E value 1.0e-116
Match length 227
% identity 99

NCBI Description Soybean mRNA for phosphoenolpyruvate carboxylase

Seq. No. 622

Contig ID 325_9.R1040 5'-most EST uaw700663617.h1

Method BLASTN
NCBI GI g218266
BLAST score 82
E value 2.0e-38
Match length 188
% identity 97

NCBI Description Glycine max mRNA for phosphoenolpyruvate carboxylase,

complete cds >gi_2175011_dbj_E06829_E06829 cDNA encoding

phosphoenolpyruvate carboxylase

>gi 2176575_dbj_E08459_E08459 cDNA encoding phosphoenol

pyruvate carboxylase

Seq. No. 623

Contig ID 328_1.R1040 5'-most EST rca701002203.h1

Method BLASTN
NCBI GI g2626748
BLAST score 180
E value 7.0e-97
Match length 180
% identity 100

NCBI Description Glycine max mRNA for phosphoenolpyruvate carboxylase,

partial cds, clone: GmPEPC4

Seq. No. 624

Contig ID 329_1.R1040 5'-most EST fde700873777.h1

Method BLASTN
NCBI GI g20906
BLAST score 146
E value 4.0e-76
Match length 506



% identity NCBI Description P.sativum mRNA for thioredoxin F isoform

Seq. No.

Contig ID

329 2.R1040

5'-most EST

LIB3138-071-P1-N1-B5

Method NCBI GI BLASTX g2266947

BLAST score E value

977

Match length % identity

1.0e-106 201 89

NCBI Description

(AF008939) phosphoenolpyruvate carboxylase 1 [Gossypium

hirsutum]

Seq. No.

626

Contig ID

329 3.R1040

5'-most EST

LIB3106-005-Q1-K2-A3

Method NCBI GI BLASTN g20906

BLAST score

78

E value Match length 9.0e-36

% identity

242 83

627

NCBI Description

P.sativum mRNA for thioredoxin F isoform

Seq. No. Contig ID

329 5.R1040

5'-most EST

ncj700987503.h1

Method NCBI GI BLASTN q2626749

BLAST score

245

E value

1.0e-135

Match length

319

% identity

94

NCBI Description

Glycine max mRNA for phosphoenolpyruvate carboxylase,

partial cds, clone: GmPEPC4

Seq. No.

628

Contig ID

329 6.R1040

5'-most EST

jC-qmle01810059a04d1

Method

BLASTN

NCBI GI

q2266946

BLAST score

92

E value

4.0e-44

Match length

132

% identity

92

NCBI Description

Gossypium hirsutum phosphoenolpyruvate carboxylase 1

(PEPC1) mRNA, complete cds

Seq. No.

629

Contig ID

331 1.R1040

5'-most EST Method

LIB3049-042-Q1-E1-A11

BLASTX

NCBI GI

q3914403

BLAST score

1640

E value

0.0e + 00



Match length % identity NCBI Description 498 65

PROTOPORPHYRINOGEN OXIDASE, MITOCHONDRIAL (PPO II) (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME II) (PPX II) >gi 2370335_emb_CAA73866_ (Y13466) protoporphyrinogen oxidase [Nicotiana tabacum] >gi 3929920 dbj BAA34712_ (AB020500) mitochondrial protoporphyrinogen oxidase

[Nicotiana tabacum] >gi_4105188 (AF044129)

protoporphyrinogen oxidase PX-2 [Nicotiana tabacum]

Seq. No.

630

Contig ID 5'-most EST 332 3.R1040 uaw700661729.h1

Seq. No.

631

Contig ID

332 4.R1040

5'-most EST

jC-gmro02800042b12a1

Method NCBI GI BLASTN g170048

BLAST score E value

156 2.0e-82

Match length

220

% identity

100

NCBI Description Glycine max SbPRP1 gene encoding a proline-rich protein,

complete cds

Seq. No.

632

Contig ID

332 5.R1040

5'-most EST

LIB3170-028-Q1-K1-F2

Seq. No.

633

Contig ID

332 6.R1040

5'-most EST

LIB3170-059-Q1-J1-H2

Seq. No.

634

Contig ID

332 7.R1040

5'-most EST

LIB3049-008-Q1-E1-E3 BLASTX

Method NCBI GI

q4490728

BLAST score

197

E value

1.0e-14

Match length

110

% identity

NCBI Description

(AL035709) putative protein [Arabidopsis thaliana]

Seq. No.

635

Contig ID

333 1.R1040

5'-most EST

LIB3092-005-Q1-K1-G11 BLASTN

Method NCBI GI

g1944318

BLAST score

875

E value

Match length

0.0e+00

% identity

1051

97 NCBI Description Glycine max mRNA for cysteine proteinase inhibitor,

complete cds



```
Seq. No.
                  636
                  333 2.R1040
Contig ID
                  jC-gmf102220077a10d1
5'-most EST
Method
                  BLASTN
                  g1944318
NCBI GI
BLAST score
                  251
E value
                  1.0e-139
Match length
                  473
% identity
                  92
NCBI Description
                  Glycine max mRNA for cysteine proteinase inhibitor,
                  complete cds
Seq. No.
                  637
Contig ID
                  333 3.R1040
                  leu701150049.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g1944341
BLAST score
                  245
E value
                  1.0e-135
Match length
                  456
% identity
                  91
NCBI Description Glycine max DNA for cysteine proteinase inhibitor, complete
                  cds
                  638
Seq. No.
Contig ID
                  333 5.R1040
5'-most EST
                  LIB3106-051-Q1-K1-E6
Method
                  BLASTN
NCBI GI
                  g1944318
BLAST score
                  151
E value
                  2.0e-79
Match length
                  339
% identity
                  90
                 Glycine max mRNA for cysteine proteinase inhibitor,
NCBI Description
                  complete cds
                  639
Seq. No.
Contig ID
                  334_1.R1040
5'-most EST
                  jC-gmro02910051h08a1
Method
                  BLASTX
NCBI GI
                  g1752734
BLAST score
                  3069
E value
                  0.0e+00
Match length
                  668
% identity
                  (D78510) beta-glucan-elicitor receptor [Glycine max]
NCBI Description
```

Seq. No. 640

335 1.R1040 Contig ID 5'-most EST sat701005149.h1

Method BLASTX NCBI GI g1408322 BLAST score 2318 E value 0.0e + 00Match length 501 % identity 91

NCBI Description (D86351) cytochrome P-450 (CYP93A2) [Glycine max]



```
641
Seq. No.
                   335 2.R1040
Contig ID
                  LIB3170-060-Q1-K1-G8
5'-most EST
                  BLASTX
Method
                   g3913192
NCBI GI
                   2659
BLAST score
E value
                   0.0e + 00
                   509
Match length
                   100
% identity
                   CYTOCHROME P450 93A1 >gi 2129824 pir S62899 cytochrome
NCBI Description
                   P450 (CYP93 A1) - soybean >gi 1232111 dbj_BAA12159_
                   (D83968) Cytochrome P-450 (CYP93A1) [Glycine max]
                   >gi 1588679_prf _2209281A cytochrome P450 [Glycine max]
                   642
Seq. No.
                   335 3.R1040
Contig ID
                   LIB3051-072-Q1-K1-A6
5'-most EST
Method
                   BLASTX
                   g3334665
NCBI GI
BLAST score
                   2452
                   0.0e + 00
E value
                   503
Match length
                   95
% identity
                   (Y10492) putative cytochrome P450 [Glycine max]
NCBI Description
                   643
Seq. No.
                   335 4.R1040
Contig ID
                   jC-gmst02400015c10d1
5'-most EST
Seq. No.
                   644
                   335 5.R1040
Contig ID
5'-most EST
                   LIB3051-054-Q1-K2-E2
                   645
Seq. No.
                   335 7.R1040
Contig ID
                   q5126796
5'-most EST
Method
                   BLASTN
                   g3334664
NCBI GI
BLAST score
                   221
E value
                   1.0e-121
Match length
                   221
% identity
                   100
                   G.max mRNA for putative cytochrome P450, clone CP5
NCBI Description
                   646
Seq. No.
                   336 1.R1040
Contig ID
                   LIB3139-056-P1-N1-H4
5'-most EST
                   BLASTX
Method
                   q1663724
NCBI GI
                   2054
BLAST score
                   0.0e + 00
E value
```

Seq. No. 647

Match length

NCBI Description

% identity

542

74

(U50846) 4-coumarate:coenzyme A ligase [Nicotiana tabacum]



```
Contig ID
                   336 3.R1040
5'-most EST
                   qsv701045038.h1
                   BLASTN
Method
NCBI GI
                   g18517
BLAST score
                   177
E value
                   9.0e-95
Match length
                   315
% identity
NCBI Description
                  Glycine max L. Merr mRNA for 4-coumarate: CoA ligase (clone
                  4CL14)
Seq. No.
                   648
                   337 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy236d07b1
Method
                  BLASTX
NCBI GI
                   g3258637
BLAST score
                   2115
E value
                   0.0e+00
Match length
                   537
                   77
% identity
                  (AF041050) 4-coumarate:CoA ligase [Populus tremuloides]
NCBI Description
Seq. No.
                   649
                   338 1.R1040
Contig ID
5'-most EST
                   jC-gmst02400078b03a1
Method
                   BLASTX
                   q3355486
NCBI GI
BLAST score
                   1358
E value
                   1.0e-150
Match length
                   419
% identity
                  (AC004218) unknown protein [Arabidopsis thaliana]
NCBI Description
                   650
Seq. No.
Contig ID
                   338_2.R1040
5'-most EST
                   g4218186_FL
Method
                   BLASTX
NCBI GI
                   q4218187
BLAST score
                   4397
E value
                   0.0e + 00
Match length
                   854
% identity
                   98
                  (AJ010165) leghemoglobin activating factor [Glycine max]
NCBI Description
                   651
Seq. No.
                   339 1.R1040
Contig ID
5'-most EST
                   jC-gmro02910013b11a1
Method
                   BLASTX
NCBI GI
                   q4468193
BLAST score
                   2258
                   0.0e+00
E value
```

Match length 502 90 % identity

NCBI Description (AJ010201) inosine monophosphate dehydrogenase [Glycine

max]

652 Seq. No.



```
339 2.R1040
Contig ID
                  ncj700980815.hl
5'-most EST
Method
                  BLASTN
                   g4468192
NCBI GI
                   208
BLAST score
                   1.0e-113
E value
                   314
Match length
                   94
% identity
NCBI Description Glycine max mRNA for inosine monophosphate dehydrogenase
Seq. No.
                   653
                   340 1.R1040
Contig ID
                   g4127861 FL
5'-most EST
                   BLASTN
Method
                   g4127861
NCBI GI
BLAST score
                   802
                   0.0e + 00
E value
Match length
                   832
                   99
% identity
                   Glycine max mRNA for glyoxalase I
NCBI Description
Seq. No.
                   654
                   340 2.R1040
Contig ID
                   rlr700898056.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g4127861
BLAST score
                   192
                   1.0e-104
E value
Match length
                   248
                   94
% identity
NCBI Description Glycine max mRNA for glyoxalase I
                   655
Seq. No.
                   341_1.R1040
Contig ID
                   jC-gmf102220132ae12a1
5'-most EST
                   BLASTX
Method
                   g4127878
NCBI GI
                   3244
BLAST score
                   0.0e + 00
E value
                   626
Match length
 % identity
                   100
                   (AJ011831) NDX1 homeobox protein [Glycine max]
NCBI Description
                   656
 Seq. No.
                   342 1.R1040
 Contig ID
                   leu701153789.h1
 5'-most EST
                   BLASTN
 Method
                   q3341442
 NCBI GI
                    1010
 BLAST score
                   0.0e + 00
 E value
 Match length
                    1023
```

NCBI Description Glycine max mRNA for root nodule acid phosphatase Seq. No.

% identity

657

100

342 2.R1040 Contig ID leu701149822.h1 5'-most EST



Method BLASTN
NCBI GI g3341442
BLAST score 219
E value 1.0e-120
Match length 386
% identity 92

NCBI Description Glycine max mRNA for root nodule acid phosphatase

Seq. No.

Contig ID 342 3.R1040

5'-most EST jC-gmfl02220114h03a1

658

Method BLASTN
NCBI GI g3341442
BLAST score 187
E value 1.0e-101
Match length 301
% identity 92

NCBI Description Glycine max mRNA for root nodule acid phosphatase

Seq. No. 659

Contig ID 343 1.R1040

5'-most EST LIB3072-047-Q1-K1-D6

Method BLASTN
NCBI GI g296442
BLAST score 621
E value 0.0e+00
Match length 703
% identity 97

NCBI Description G.max ADR11 mRNA

Seq. No. 660

Contig ID 343_2.R1040

5'-most EST LIB3028-010-Q1-B1-C1

Method BLASTN
NCBI GI g296442
BLAST score 376
E value 0.0e+00
Match length 721
% identity 94

NCBI Description G.max ADR11 mRNA

Seq. No. 661

Contig ID 343 3.R1040

5'-most EST LIB3170-024-Q1-K1-B4

Method BLASTN
NCBI GI g296442
BLAST score 101
E value 3.0e-49
Match length 288
% identity 93

NCBI Description G.max ADR11 mRNA

Seq. No. 662

Contig ID 344 3.R1040 5'-most EST g5678078 Method BLASTN NCBI GI g296444

Mary .



BLAST score 238
E value 1.0e-131
Match length 242
% identity 100
NCBI Description G.max ADR6 mRNA

Seq. No. 663

Contig ID 344_4.R1040

5'-most EST jC-gmro02910013g10a1

Method BLASTN
NCBI GI g296444
BLAST score 170
E value 1.0e-90
Match length 182
% identity 99

NCBI Description G.max ADR6 mRNA

Seq. No. 664

Contig ID 344_5.R1040

5'-most EST LIB3106-102-Q1-K1-C11

Method BLASTN
NCBI GI g2317899
BLAST score 172
E value 1.0e-91
Match length 184
% identity 98

NCBI Description Glycine max Sali3-2 mRNA, complete cds

Seq. No. 665

Contig ID 344_6.R1040

5'-most EST uC-gmropic061g09b1

Method BLASTX
NCBI GI g2317900
BLAST score 930
E value 1.0e-100
Match length 279
% identity 62

NCBI Description (U89693) Sali3-2 [Glycine max]

Seq. No. 666

344 7.R1040 Contig ID g4290365 5'-most EST BLASTN Method g2317899 NCBI GI BLAST score 523 0.0e+00E value 1054 Match length 91 % identity

NCBI Description Glycine max Sali3-2 mRNA, complete cds

Seq. No. 667

Contig ID 344_8.R1040 5'-most EST jsh701069031.h1

Method BLASTN
NCBI GI g296444
BLAST score 649
E value 0.0e+00



Match length % identity

NCBI Description G.max ADR6 mRNA

668 Seq. No.

346 1.R1040 Contig ID bth700848391.h1 5'-most EST

BLASTN Method q2687725 NCBI GI 12-32 BLAST score 0.0e + 00E value 1357 Match length % identity

NCBI Description Glycine max mRNA for putative 2-hydroxydihydrodaidzein

reductase

669 Seq. No.

346 3.R1040 Contig ID

jC-gmro02800032e12d1 5'-most EST

BLASTN Method q2687725 NCBI GI BLAST score 172 7.0e-92 E value 354 Match length % identity

Glycine max mRNA for putative 2-hydroxydihydrodaidzein NCBI Description

reductase

670 Seq. No.

348 1.R1040 Contig ID

LIB3093-001-Q1-K1-B4 5'-most EST

BLASTX Method q3334449 NCBI GI 1703 BLAST score 0.0e+00E value 321 Match length % identity 100

ALTERNATIVE OXIDASE 1 PRECURSOR >gi 395216 emb CAA48653 NCBI Description

(X68702) alternative oxidase [Glycine max]

>gi_740229_prf__2004454A respiratory chain terminal oxidase

[Glycine max]

671 Seq. No.

350 1.R1040 Contig ID vwf700677744.hl 5'-most EST

BLASTN Method g479144 NCBI GI 553 BLAST score 0.0e + 00E value 611 Match length 97 % identity

NCBI Description G.max mRNA for ATP synthase subunit

672 Seq. No.

351 1.R1040 Contig ID

5'-most EST LIB3072-052-Q1-E1-E12

BLASTN Method



```
g18540
NCBI GI
                   432
BLAST score
                   0.0e + 00
E value
                   481
Match length
                   98
% identity
NCBI Description G.max BBI mRNA for proteinase inhibitor
                   673
Seq. No.
                   351 2.R1040
Contig ID
                   LIB3029-009-Q1-B1-A6
5'-most EST
                   BLASTN
Method
                   g18540
NCBI GI
                   200
BLAST score
                   1.0e-108
E value
                   431
Match length
                   95
% identity
NCBI Description G.max BBI mRNA for proteinase inhibitor
                   674
Seq. No.
                   352 1.R1040
Contig ID
                   euj700697942.h1
5'-most EST
                   BLASTX
Method
                   g434061
NCBI GI
                   2023
BLAST score
                   0.0e + 00
E value
                   421
Match length
                   91
% identity
NCBI Description (D16107) basic 7S globulin [Glycine max]
                   675
Seq. No.
                   352 2.R1040
Contig ID
                   zpv700757208.h1
5'-most EST
                   BLASTN
Method
                   q434060
NCBI GI
                   347
BLAST score
                   0.0e + 00
E value
                   483
Match length
                   95
% identity
NCBI Description Soybean DNA for basic 7S globulin, complete cds
                   676
Seq. No.
                   352 6.R1040
Contig ID
                   txt700732264.h1
 5'-most EST
                   BLASTN
Method
NCBI GI
                   q18542
                   220
BLAST score
                   1.0e-120
E value
                    343
Match length
                    98
 % identity
                   Soybean Bg gene for basic 7S globulin
NCBI Description
                    677
 Seq. No.
```

.. . -

Contig ID 353_1.R1040
5'-most EST g3334658_FL
Method BLASTX
NCBI GI g3334659
BLAST score 2330



E value 0.0e+00 Match length 480 95

NCBI Description (Y10489) putative cytochrome P450 [Glycine max]

Seq. No. 678

354 1.R1040 Contig ID 5'-most EST g3334660_FL Method BLASTX NCBI GI g3334661 BLAST score 2337 E value 0.0e + 00Match length 496 % identity 92

NCBI Description (Y10490) putative cytochrome P450 [Glycine max]

Seq. No. 679

Contig ID 354 3.R1040 5'-most EST leu701150560.h1 Method BLASTX NCBI GI g3334661

BLAST score 526 E value 1.0e-53 Match length 154 % identity 69

NCBI Description (Y10490) putative cytochrome P450 [Glycine max]

Seq. No. 680

Contig ID 355 1.R1040 5'-most EST g3334662 FL Method BLASTX NCBI GI g3334663 BLAST score 2601 E value 0.0e+00Match length 522 % identity 97

NCBI Description (Y10491) putative cytochrome P450 [Glycine max]

Seq. No. 681

Contig ID 355 2.R1040

5'-most EST LIB3170-032-Q1-J1-E8

Method BLASTX
NCBI GI g2765093
BLAST score 229
E value 2.0e-20
Match length 74
% identity 68

NCBI Description (Y10983) putative cytochrome P450 [Glycine max]

Seq. No. 682

Contig ID 357_1.R1040
5'-most EST g3334666_FL
Method BLASTX
NCBI GI g3334667
BLAST score 2548
E value 0.0e+00
Match length 504



% identity (Y10493) putative cytochrome P450 [Glycine max] NCBI Description

Seq. No. 683

358 1.R1040 Contig ID

jC-gmro02910006h12a1 5'-most EST

BLASTX Method q398957 NCBI GI BLAST score 2529 0.0e + 00E value 484 Match length 99 % identity

1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE (ACC SYNTHASE) NCBI Description

(S-ADENOSYL-L-METHIONINE METHYLTHIOADENOSINE-LYASE)

>gi_99885_pir__S25002 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) - soybean >gi_18558_emb_CAA47474_ (X67100) 1-aminocyclopropane 1-carboxylate synthase

[Glycine max] >gi_1090777_prf__2019442A aminocyclopropane

carboxylate synthase [Glycine max]

Seq. No. 684

359 1.R1040 Contig ID

LIB3055-002-Q1-B1-F10 5'-most EST

BLASTX Method g3059095 NCBI GI 6994 BLAST score 0.0e + 00E value 1383 Match length % identity

(AJ001091) magnesium chelatase subunit [Glycine max] NCBI Description

685 Seq. No.

359 2.R1040 Contig ID

5'-most EST fC-gmle7000743672r1

BLASTN Method NCBI GI q3059094 BLAST score 204 1.0e-111 E value 324 Match length % identity 91

Glycine max mRNA for magnesium chelatase subunit NCBI Description

Seq. No. 686

359 3.R1040 Contig ID

5'-most EST fC-gmle700557852f2

BLASTN Method NCBI GI g3059094 BLAST score 164 3.0e-87 E value Match length 212 % identity

Glycine max mRNA for magnesium chelatase subunit NCBI Description

687 Seq. No.

360 1.R1040 Contig ID

LIB3072-010-Q1-E1-B3 5'-most EST

BLASTN Method



```
NCBI GI g169944
BLAST score 425
E value 0.0e+00
Match length 476
% identity 98
NCBI Description Soybean
```

NCBI Description Soybean protease inhibitor C-II mRNA, complete cds

Seq. No. 688

Contig ID 361_1.R1040

5'-most EST jC-gmro02910036a01a1

Method BLASTX
NCBI GI g3915111
BLAST score 2471
E value 0.0e+00
Match length 506
% identity 95

NCBI Description TRANS-CINNAMATE 4-MONOOXYGENASE (CINNAMIC ACID

4-HYDROXYLASE) (CA4H) (P450C4H) (CYTOCHROME P450 73) >qi 1044868 emb CAA63172 (X92437) cinnamic acid

4-hydroxylase [Glycine max]

Seq. No. 689

361 2.R1040 Contig ID $q42\overline{9}8689$ 5'-most EST BLASTN Method q1044867 NCBI GI BLAST score 287 1.0e-160 E value 454 Match length 94 % identity

NCBI Description Glycine max mRNA for cinnamic acid 4-hydroxylase (CYP73)

Seq. No. 690

Contig ID 361_3.R1040 5'-most EST taw700660393.h1

Method BLASTN
NCBI GI g1044867
BLAST score 248
E value 1.0e-137
Match length 443
% identity 90

NCBI Description Glycine max mRNA for cinnamic acid 4-hydroxylase (CYP73)

Seq. No. 691

Contig ID 362 1.R1040 5'-most EST jsh701065321.h1

Method BLASTX
NCBI GI g1362047
BLAST score 1962
E value 0.0e+00
Match length 380
% identity 97

NCBI Description cysteine proteinase (EC 3.4.22.-) precursor - soybean

>gi_479060_emb_CAA83673_ (Z32795) cysteine proteinase
[Glycine max] >gi_1096153_prf__2111244A Cys protease

[Glycine max]



```
Seq. No.
Contig ID
                  362 8.R1040
                  awf700839974.hl
5'-most EST
                  BLASTN
Method
                  q479059
NCBI GI
BLAST score
                  126
                  1.0e-64
E value
                  138
Match length
% identity
NCBI Description G.max (Fiskeby V) mRNA for cysteine endopeptidase
Seq. No.
                   363 1.R1040
Contig ID
                  LIB3072-023-Q1-E1-B2
5'-most EST
                  BLASTN
Method
                   q18571
NCBI GI
BLAST score
                   416
                   0.0e + 00
E value
                   444
Match length
                   99
% identity
NCBI Description G.max D-II mRNA for proteinase isoinhibitor D-II
                   694
Seq. No.
                   365 1.R1040
Contig ID
                   LIB3073-008-Q1-K1-E7
5'-most EST
                   BLASTN
Method
NCBI GI
                   q886869
                   659
BLAST score
                   0.0e+00
E value
                   789
Match length
                   96
% identity
                   G.max ENOD40-2 gene
NCBI Description
                   695
Seq. No.
                   367 1.R1040
Contig ID
                   g218253 FL
5'-most EST
                   BLASTN
Method
                   g218253
NCBI GI
                   643
BLAST score
                   0.0e + 00
E value
                   735
Match length
                   100
% identity
NCBI Description Soybean mRNA for early nodulin
                   696
Seq. No.
                   367 2.R1040
Contig ID
                   LIB3170-042-Q1-J1-A10
5'-most EST
                   BLASTX
Method
                   q134892
NCBI GI
BLAST score
                   357
                   1.0e-33
```

E value Match length 146 46 % identity

SIGNAL RECOGNITION PARTICLE RECEPTOR ALPHA SUBUNIT NCBI Description

(SR-ALPHA) (DOCKING PROTEIN ALPHA) (DP-ALPHA)

>gi_88607_pir__A29440 signal recognition particle receptor - human $>\overline{gi}$ 30866 emb CAA29608_ (X06272) docking protein



[Homo sapiens] >gi_4507223_ref_NP_003130.1_pSRPR_ signal recognition particle receptor ('docking protein')

Seq. No. 697

Contig ID 367_3.R1040

5'-most EST LIB3139-078-P1-N1-A12

Method BLASTX
NCBI GI g134892
BLAST score 184
E value 4.0e-13
Match length 125
% identity 42

NCBI Description SIGNAL RECOGNITION PARTICLE RECEPTOR ALPHA SUBUNIT

(SR-ALPHA) (DOCKING PROTEIN ALPHA) (DP-ALPHA)

>gi_88607_pir_A29440 signal recognition particle receptor - human >gi_30866_emb_CAA29608_ (X06272) docking protein [Homo sapiens] >gi_4507223_ref_NP_003130.1_psrPR_ signal

recognition particle receptor ('docking protein')

Seq. No. 698

Contig ID 367_4.R1040 5'-most EST rlr700898123.h1

Method BLASTX
NCBI GI g134891
BLAST score 392
E value 8.0e-38
Match length 113
% identity 66

NCBI Description SIGNAL RECOGNITION PARTICLE RECEPTOR ALPHA SUBUNIT

(SR-ALPHA) (DOCKING PROTEIN ALPHA) (DP-ALPHA)

>gi_89066_pir_A24570 signal recognition particle receptor

precursor - dog >gi_997_emb_CAA26945_ (X03184) signal

recognition particle receptor [Canis sp.]

>gi_224778_prf__1112224A signal recognition particle

receptor [Canis familiaris]

Seq. No. 699

Contig ID 367_5.R1040

5'-most EST LIB3092-036-Q1-K1-F9

Method BLASTN
NCBI GI g18584
BLAST score 316
E value 1.0e-177
Match length 562
% identity 98

NCBI Description G.max ENOD55-1 mRNA

Seq. No. 700

Contig ID 370_1.R1040

5'-most EST LIB3051-027-Q1-K1-D6

Method BLASTN
NCBI GI g2052028
BLAST score 824
E value 0.0e+00
Match length 909
% identity 98

NCBI Description G.max mRNA for glutathione transferase



```
701
Seq. No.
Contig ID
                   370 2.R1040
5'-most EST
                   jC-gmle01810024d08a1
Method
                   BLASTN
NCBI GI
                   q2920665
BLAST score
                   779
E value
                   0.0e+00
Match length
                   858
% identity
                   97
                  Glycine max 2,4-D inducible glutathione S-transferase
NCBI Description
                   (GSTa) mRNA, complete cds
                   702
Seq. No.
Contig ID
                   370 3.R1040
5'-most EST
                   LIB3051-051-Q1-K1-A5
                   BLASTN
Method
NCBI GI
                   g2920665
BLAST score
                   329
E value
                   0.0e + 00
Match length
                   341
% identity
                   99
                   Glycine max 2,4-D inducible glutathione S-transferase
NCBI Description
                   (GSTa) mRNA, complete cds
Seq. No.
                   703
Contig ID
                   370 5.R1040
5'-most EST
                   LIB3051-097-Q1-K1-E8
Method
                   BLASTN
                   g2920665
NCBI GI
BLAST score
                   252
                   1.0e-139
E value
                   284
Match length
                   97
% identity
                   Glycine max 2,4-D inducible glutathione S-transferase
NCBI Description
                   (GSTa) mRNA, complete cds
                   704
Seq. No.
                   374 1.R1040
Contig ID
                   pxt700943646.hl
5'-most EST
                   BLASTN
Method
                   g505584
NCBI GI
BLAST score
                   915
E value
                   0.0e + 00
Match length
                   931
% identity
                   100
NCBI Description
                  G.max mRNA for Glyoxalase I
```

705 Seq. No. 376 1.R1040 Contig ID

5'-most EST fC-gmst700653759a1 BLASTX Method NCBI GI a121336 BLAST score 1799 E value 0.0e + 00Match length 342 % identity 98



NCBI Description GLUTAMINE SYNTHETASE CYTOSOLIC ISOZYME (GLUTAMATE--AMMONIA LIGASE) >gi_81774_pir__JQ0937 glutamate--ammonia ligase (EC 6.3.1.2), cytosolic - soybean >gi_256143_bbs_113802

(S46513) cytosolic glutamine synthetase, GS {EC 6.3.1.2} [Glycine max=soybeans, var Prize, Peptide, 355 aa] [Glycine

max]

Seq. No. 706

Contig ID 376_2.R1040 5'-most EST asj700967306.h1

Method BLASTX
NCBI GI g3128203
BLAST score 465
E value 3.0e-61
Match length 171
% identity 73

NCBI Description (AC004521) unknown protein [Arabidopsis thaliana]

Seq. No. 707

Contig ID 376 3.R1040

5'-most EST LIB3109-056-Q1-K1-D5

Method BLASTN
NCBI GI g547507
BLAST score 1332
E value 0.0e+00
Match length 1431
% identity 99

NCBI Description G.max mRNA for glutamine synthetase

Seq. No. 708

Contig ID 376_4.R1040
5'-most EST g587119_FL
Method BLASTX
NCBI GI g587120
BLAST score 1756

E value 0.0e+00
Match length 340
% identity 96

NCBI Description (X81460) glutamate--ammonia ligase [Glycine max]

Seq. No. 709

Contig ID 376 5.R1040

5'-most EST jC-gmst02400005f10d1

Method BLASTN
NCBI GI g256142
BLAST score 397
E value 0.0e+00
Match length 458
% identity 95

NCBI Description cytosolic glutamine synthetase [Glycine max=soybeans, var

Prize, mRNA, 1450 nt]

Seq. No. 710

Contig ID 376 6.R1040

5'-most EST fC-gmst700653759h1

Method BLASTN NCBI GI g256142

436



```
BLAST score
                    453
E value
                    0.0e + 00
Match length
                    707
% identity
                    95
```

NCBI Description cytosolic glutamine synthetase [Glycine max=soybeans, var Prize, mRNA, 1450 nt]

Seq. No. 711

Contig ID 376_7.R1040 5'-most EST pcp700994236.h1

Seq. No.

712 Contig ID 376 10.R1040 5'-most EST zhf700959375.h1

Seq. No.

Contig ID 376 11.R1040 5'-most EST fde700875135.h1

713

Method BLASTN NCBI GI g312300 BLAST score 77 E value 3.0e-35 Match length 242 % identity 94

NCBI Description G.max cDNA for glutamine synthetase (3' region)

Seq. No. 714

Contig ID 379 1.R1040

5'-most EST LIB3051-029-Q1-K1-E5

Method BLASTN NCBI GI q18642 BLAST score 720 E value 0.0e+00Match length 737 % identity 99

NCBI Description G.max mRNA from stress-induced gene (H4)

Seq. No. 715

Contig ID 379 2.R1040

5'-most EST LIB3051-111-Q1-K1-C1

Method BLASTN NCBI GI g18743 BLAST score 743 E value 0.0e+00Match length 771 % identity 99

NCBI Description G.max mRNA from stress-induced gene (SAM22)

Seq. No.

716

Contig ID 379 4.R1040

5'-most EST LIB3051-111-Q1-K1-G5

Method BLASTN NCBI GI g18642 BLAST score 409 E value 0.0e+00Match length 662 % identity 93





NCBI Description G.max mRNA from stress-induced gene (H4)

717 Seq. No. Contig ID 379 5.R1040 5'-most EST g4396357 Method BLASTN NCBI GI g18743 BLAST score 266 E value 1.0e-148 Match length 478 % identity 93

NCBI Description G.max mRNA from stress-induced gene (SAM22)

Contig ID 379_6.R1040
5'-most EST zsg701125214.h1
Method BLASTN
NCBI GI g18743
BLAST score 425
E value 0.0e+00
Match length 547

Seq. No.

% identity

NCBI Description G.max mRNA from stress-induced gene (SAM22)

Seq. No. 719

Contig ID 379_10.R1040

5'-most EST LIB3051-048-Q1-K1-C11

96

718

Method BLASTN
NCBI GI g18743
BLAST score 337
E value 0.0e+00
Match length 431
% identity 98

NCBI Description G.max mRNA from stress-induced gene (SAM22)

Seq. No. 720

Contig ID 379 15.R1040

5'-most EST LIB3051-022-Q1-K1-F2

Method BLASTN
NCBI GI g18743
BLAST score 140
E value 9.0e-73
Match length 309
% identity 93

NCBI Description G.max mRNA from stress-induced gene (SAM22)

Seq. No. 721

Contig ID 380_1.R1040 5'-most EST kll701212864.h1

Method BLASTN
NCBI GI g18644
BLAST score 761
E value 0.0e+00
Match length 761
% identity 100

NCBI Description Soybean mRNA for HMG-1 like protein

```
Seq. No.
 Contig ID
                    380 2.R1040
 5'-most EST
                    LIB3167-078-P1-K2-H1
 Method
                    BLASTN
 NCBI GI
                    g18644
 BLAST score
                    148
 E value
                    2.0e-77
 Match length
                    172
 % identity
                    97
 NCBI Description Soybean mRNA for HMG-1 like protein
 Seq. No.
                    723
 Contig ID
                    380 3.R1040
 5'-most EST
                    jC-gmst02400001g11a1
 Method
                    BLASTN
 NCBI GI
                    q18644
 BLAST score
                    287
 E value
                    1.0e-160
 Match length
                    339
 % identity
                    96
NCBI Description
                  Soybean mRNA for HMG-1 like protein
 Seq. No.
                   724
Contig ID
                   380 4.R1040
 5'-most EST
                   uC-gmflminsoy069c08b1
Method
                   BLASTN
NCBI GI
                   g18644
BLAST score
                   148
E value
                   2.0e-77
Match length
                   311
 % identity
                   95
NCBI Description
                   Soybean mRNA for HMG-1 like protein
Seq. No.
                   725
                   381 1.R1040
Contig ID
5'-most EST
                   kl1701213229.h1
Method
                   BLASTN
NCBI GI
                   g18646
BLAST score
                   791
E value
                   0.0e+00
Match length
                   832
```

Seq. No. 726

% identity

Contig ID 381 2.R1040

5'-most EST LIB3039-026-Q1-E1-H1

99

Method BLASTN
NCBI GI g18648
BLAST score 577
E value 0.0e+00

Match length 585 % identity 100

NCBI Description Soybean mRNA for HMG-Y related protein, variant B

NCBI Description Soybean mRNA for HMG-Y related protein, variant A

Seq. No. 727

Contig ID 383_1.R1040



```
5'-most EST
                   g662923 FL
Method
                   BLASTN
NCBI GI
                   g662923
BLAST score
                   504
E value
                   0.0e + 00
Match length
                   654
                   100
% identity
NCBI Description
                  G.max mRNA for heat shock transcription factor
Seq. No.
                   728
Contig ID
                   384 1.R1040
5'-most EST
                   jC-gmro02910010e01a1
                   BLASTN
```

Method BLASTN
NCBI GI g662925
BLAST score 1373
E value 0.0e+00
Match length 1373
% identity 100

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 729
Contig ID 384_2.R1040
5'-most EST jC-gmro02910074f07a1
Method BLASTN

NCBI GI g662927
BLAST score 1188
E value 0.0e+00
Match length 1188
% identity 100

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 730

Contig ID 384_3.R1040 5'-most EST ncj700982933.h1

Seq. No. 731

Contig ID 385 1.R1040 5'-most EST g662931 FL Method BLASTX NCBI GI g2129832 BLAST score 1659 E value 0.0e+00Match length 370 % identity 87

NCBI Description heat shock transcription factor 5 - soybean

>gi_671868_emb_CAA87080_ (Z46956) heat shock transcription

factor 5 [Glycine max]

Seq. No. 732 Contig ID 385 2.R1040 5'-most EST g662926 FL Method BLASTN NCBI GI g662926 BLAST score 988 E value 0.0e + 00Match length 1014 % identity 100

204





NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 733

Contig ID 385_3.R1040

5'-most EST fC-gmse700725253a1

Method BLASTN
NCBI GI g662931
BLAST score 212
E value 1.0e-115
Match length 489
% identity 89

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 734

Contig ID 387_1.R1040 5'-most EST bth700845329.h1

Method BLASTN
NCBI GI g662929
BLAST score 1376
E value 0.0e+00
Match length 1380
% identity 100

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 735

Contig ID 387 2.R1040 5'-most EST g5606125 Method BLASTN NCBI GI g662929 BLAST score 465 E value 0.0e+00Match length 541 % identity 97

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 736

Contig ID 387 3.R1040 5'-most EST dpv701100361.h1

Method BLASTN
NCBI GI g662929
BLAST score 161
E value 3.0e-85
Match length 427
% identity 90

NCBI Description G.max mRNA for heat shock transcription factor

Seq. No. 737

Contig ID 387 4.R1040

5'-most EST uC-gmrominsoy044a01b1

Method BLASTN
NCBI GI g662929
BLAST score 281
E value 1.0e-157
Match length 472
% identity 94

NCBI Description G.max mRNA for heat shock transcription factor



Seq. No. 738

Contig ID 389_1.R1040 5'-most EST epx701104693.h1

Method BLASTN
NCBI GI g18658
BLAST score 547
E value 0.0e+00
Match length 668
% identity 99

NCBI Description Soybean hsp22 mRNA

Seq. No. 739

Contig ID 390_1.R1040 5'-most EST hrw701060990.h1

Method BLASTX
NCBI GI g418854
BLAST score 1149
E value 1.0e-126
Match length 230
% identity 25

NCBI Description ubiquitin precursor - parsley >gi_288112_emb_CAA45621_

(X64344) polyubiquitin [Petroselinum crispum] >gi_288114_emb_CAA45622_ (X64345) polyubiquitin

[Petroselinum crispum]

Seq. No. 740

Contig ID 390_2.R1040

5'-most EST LIB3028-001-Q1-B1-C10

Method BLASTX
NCBI GI g418854
BLAST score 1528
E value 1.0e-170
Match length 306
% identity 33

NCBI Description ubiquitin precursor - parsley >gi_288112_emb_CAA45621

(X64344) polyubiquitin [Petroselinum crispum] >gi_288114_emb_CAA45622_ (X64345) polyubiquitin

[Petroselinum crispum]

Seq. No. 741

Contig ID 391_1.R1040 5'-most EST epx701105962.h1

Method BLASTN
NCBI GI g1575724
BLAST score 712
E value 0.0e+00
Match length 988
% identity 98

NCBI Description Glycine max 14-3-3 related protein SGF14A mRNA, complete

cds

Seq. No. 742

Contig ID 391 2.R1040

5'-most EST LIB3051-029-Q1-K1-G4

Method BLASTN
NCBI GI g1575724
BLAST score 202



E value 1.0e-109 Match length 426 % identity 91

NCBI Description Glycine max 14-3-3 related protein SGF14A mRNA, complete

cds

Seq. No. 743

Contig ID 392_1.R1040

5'-most EST uC-gmropic050e09b1

Method BLASTN
NCBI GI g3452090
BLAST score 390
E value 0.0e+00
Match length 398
% identity 99

NCBI Description Glycine max mRNA for rudimentary enhancer, partial

Seq. No. 744

Contig ID 392_2.R1040 5'-most EST kll701213706.h1

Method BLASTN
NCBI GI g3452090
BLAST score 202
E value 1.0e-110
Match length 243
% identity 96

NCBI Description Glycine max mRNA for rudimentary enhancer, partial

Seq. No. 745

Contig ID 393_1.R1040

5'-most EST LIB3051-007-Q1-E1-D10

Method BLASTN
NCBI GI g3452136
BLAST score 235
E value 1.0e-129
Match length 243
% identity 100

NCBI Description Glycine max mRNA for glucose-6-phosphate-dehydrogenase,

partial

Seq. No. 746

Contig ID 393_2.R1040 5'-most EST hrw701063583.h1

Method BLASTN
NCBI GI g603218
BLAST score 392
E value 0.0e+00
Match length 752
% identity 88

NCBI Description Medicago sativa glucose-6-phosphate dehydrogenase mRNA,

complete cds

Seq. No. 747

Contig ID 395_1.R1040 5'-most EST rlr700896427.h1

Method BLASTN NCBI GI g3452143



```
BLAST score
E value
                   0.0e + 00
Match length
                   643
% identity
                   96
NCBI Description Glycine max mRNA for chalcone isomerase, partial
                   748
Seq. No.
                   396 1.R1040
Contig ID
5'-most EST
                   LIB3051-107-Q1-K1-D2
Method
                  BLASTN
NCBI GI
                   g510514
BLAST score
                   683
E value
                   0.0e+00
Match length
                   747
% identity
                   98
NCBI Description
                  G.max KTi-S mRNA
                   749
Seq. No.
                   397 1.R1040
Contig ID
5'-most EST
                  LIB3065-009-Q1-N1-C8
Method
                  BLASTX
NCBI GI
                  g2129827
BLAST score
                   2163
E value
                  0.0e+00
Match length
                   473
% identity
                   91
NCBI Description
                  embryonic abundant protein group 3 precursor (clone PM10) -
                   soybean >gi 414977 (U02966) 51 kDa seed maturation protein
                   [Glycine max]
Seq. No.
                  750
Contig ID
                  397 2.R1040
5'-most EST
                  g311697 FL
Method
                  BLASTX
NCBI GI
                  g421875
BLAST score
                  1921
E value
                  0.0e + 00
Match length
                  458
% identity
                  83
NCBI Description
                  Lea protein - soybean >gi 311698 emb CAA80491 (Z22872) Lea
                  protein [Glycine max]
Seq. No.
                  751
                  398 1.R1040
Contig ID
5'-most EST
                  zpv700762262.h1
Method
                  BLASTX
NCBI GI
                  g505138
BLAST score
                  4431
E value
                  0.0e + 00
Match length
                  866
```

% identity

(D13949) lipoxygenase-2 [Glycine max] NCBI Description

Seq. No. 752

398 2.R1040 Contig ID 5'-most EST q18674 FL Method BLASTX



```
NCBI GI
                   q126398
BLAST score
                   4277
                   0.0e+00
E value
Match length
                   839
                   98
% identity
                   SEED LIPOXYGENASE-1 (L-1) >gi 2144382 pir DASYL2
NCBI Description
                   lipoxygenase (EC 1.13.11.12) \overline{1} - cybean
                   >gi 1000060 pdb 2SBL B Lipoxygenase-1 (Soybean)
                   (E.C.1.13.11.12) >gi 2392632 pdb 1YGE
                                                            Lipoxygenase-1
                   (Soybean) At 100k >gi_18675_emb_CAA47717_ (X67304)
                   lipoxygenase [Glycine max] \overline{>}gi \overline{2}95118 (J\overline{0}2795)
                   lipoxygenase-1 [Glycine max]
                   753
Seq. No.
Contig ID
                   400 1.R1040
5'-most EST
                   uC-gmrominsoy125f04b1
Method
                   BLASTN
NCBI GI
                   g1262439
BLAST score
                   229
E value
                   1.0e-125
                   585
Match length
% identity
                   85
NCBI Description Glycine max lipoxygenase (vlxC) mRNA, complete cds
Seq. No.
                   754
                   400 2.R1040
Contig ID
5'-most EST
                   dpv701096996.h1
                   BLASTX
Method
                   q436169
NCBI GI
BLAST score
                   752
                   6.0e-80
E value
Match length
                   219
% identity
                   64
NCBI Description
                  (U04526) lipoxygenase [Glycine max]
                   755
Seq. No.
                   400 3.R1040
Contig ID
                   dpv701102162.h1
5'-most EST
Method
                   BLASTX
                   q3319774
NCBI GI
BLAST score
                   174
E value
                   3.0e-12
Match length
                   43
                   72
% identity
NCBI Description
                  (Y16228) TOM7 protein [Solanum tuberosum]
Seq. No.
                   756
Contig ID
                   400 4.R1040
                   fC-gmse7000756638a1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g126409
BLAST score
                   1062
```

E value 0.0e+00Match length 739 % identity 78

LIPOXYGENASE 1 >gi 100005 pir S22153 lipoxygenase (EC NCBI Description 1.13.11.12) - kidney bean >gi 21017 emb CAA45088 (X63525)



lipoxygenase [Phaseolus vulgaris]

```
Seq. No.
                   757
Contig ID
                   400 5.R1040
5'-most EST
                   uC-gmflminsoy016g08b1
Method
                   BLASTX
NCBI GI
                   g1262440
BLAST score
                   1354
E value
                   0.0e+00
Match length
                   861
% identity
                   94
NCBI Description
                   (U26457) lipoxygenase [Glycine max]
Seq. No.
                   758
                   400 6.R1040
Contig ID
5'-most EST
                   uC-gmflminsoy011c10b1
Method
                   BLASTX
NCBI GI
                   g126411
BLAST score
                   1189
E value
                   0.0e + 00
Match length
                   874
                   61
% identity
                   SEED LIPOXYGENASE >gi_99932_pir__S13381 lipoxygenase (EC
NCBI Description
                   1.13.11.12) - soybean >gi_18746_emb_CAA39604_ (X56139)
                   lipoxygenase [Glycine max]
                   759
Seq. No.
                   400_7.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy256a03b1
Method
                   BLASTX
NCBI GI
                   g126404
BLAST score
                   1058
E value
                   0.0e+00
Match length
                   873
% identity
                   49
                   SEED LIPOXYGENASE-2 (L-2) >gi_66331_pir__DASYL1 lipoxygenase (EC 1.13.11.12) 2 - soybean >gi_170014
NCBI Description
                   (J03211) lipoxygenase (EC 1.13.11.12) [Glycine max]
                   760
Seq. No.
                   401 1.R1040
Contig ID
5'-most EST
                   ncj700981157.h1
Method
                   BLASTX
NCBI GI
                   g116157
BLAST score
                   1901
E value
                   0.0e+00
Match length
                   455
% identity
                   84
NCBI Description
                   G2/MITOTIC-SPECIFIC CYCLIN S13-6 (B-LIKE CYCLIN)
                   >gi_99895_pir__S16522 cyclin, mitotic - soybean
                   >gi 18682_emb_CAA44632_ (X62820) mitotic cyclin [Glycine
                   max]
```

761 Seq. No.

Contig ID 403 1.R1040

5'-most EST uC-gmropic091h10b1

Method BLASTN



NCBI GI g18688
BLAST score 497
E value 0.0e+00
Match length 813
% identity 96

NCBI Description Soybean nodulin-20 gene

Seq. No. 762

404 1.R1040 Contig ID 5'-most EST g18693 FL Method BLASTN NCBI GI g18693 BLAST score 810 E value 0.0e + 00810 Match length 100 % identity

NCBI Description Soybean mRNA for Nodulin-21

Seq. No. 763

Contig ID 406_1.R1040 5'-most EST kl1701203084.h1

Method BLASTN
NCBI GI g18709
BLAST score 1060
E value 0.0e+00
Match length 1107
% identity 99

NCBI Description Soybean mRNA for nodulin-26

Seq. No. 764

409 1.R1040 Contig ID 5'-most EST g18711 FL Method BLASTN NCBI GI g18711 BLAST score 892 E value 0.0e+00Match length 892 % identity 100

NCBI Description Soybean mRNA for nodulin-27

Seq. No. 765

410 1.R1040 Contig ID 5'-most EST g18717 FL Method BLASTN NCBI GI q18717 BLAST score 1299 E value 0.0e+001303 Match length 100 % identity

NCBI Description Soy bean mRNA for nodulin (E27)

Seq. No. 766

Contig ID 410_2.R1040 5'-most EST g18715_FL Method BLASTN NCBI GI g18715 BLAST score 954

211



```
E value 0.0e+00
Match length 954
% identity 100
NCBI Description Soy bean mRNA for nodulin (C51)
Seq. No. 767
Contig ID 415_1.R1040
5'-most EST g2765090 FL
```

Contig ID 415_1.R1040
5'-most EST g2765090_FI
Method BLASTX
NCBI GI g2765091
BLAST score 2601
E value 0.0e+00
Match length 527
% identity 95

NCBI Description (Y10982) putative cytochrome P450 [Glycine max]

Contig ID 416_1.R1040 5'-most EST uC-gmropic050h11b1 Method BLASTX NCBI GI g2765093 BLAST score 1441

Seq. No.

NCBI GI g2765093
BLAST score 1441
E value 0.0e+00
Match length 527
% identity 91

NCBI Description (Y10983) putative cytochrome P450 [Glycine max]

Seq. No. 769

Contig ID 416_2.R1040

5'-most EST uC-gmflminsoy040f06b1

768

Method BLASTX
NCBI GI g2765093
BLAST score 541
E value 1.0e-105
Match length 293
% identity 69

NCBI Description (Y10983) putative cytochrome P450 [Glycine max]

Seq. No. 770

Contig ID 416 3.R1040

Method BLASTN
NCBI GI g2765092
BLAST score 86
E value 2.0e-40
Match length 198
% identity 86

NCBI Description Glycine max mRNA for cytochrome P450-like protein, clone

CP9

Seq. No. 771

Contig ID 417_1.R1040 5'-most EST leu701144651.h1

Method BLASTN
NCBI GI g18723
BLAST score 1082
E value 0.0e+00



Match length 1109 % identity 99

NCBI Description Glycine max mRNA for pyrroline-5-carboxylate reductase

Seq. No. 772

Contig ID 418 1.R1040

5'-most EST LIB3040-022-Q1-E1-A7

Method BLASTN
NCBI GI g18725
BLAST score 783
E value 0.0e+00
Match length 874
% identity 97

NCBI Description Soybean (G. max) mRNA for proliferating cell nuclear

antigen (PCNA), partial

Seq. No. 773

Contig ID 418 2.R1040

5'-most EST LIB3040-005-Q1-E1-E6

Method BLASTN
NCBI GI g18725
BLAST score 311
E value 1.0e-175
Match length 315
% identity 100

NCBI Description Soybean (G. max) mRNA for proliferating cell nuclear

antigen (PCNA), partial

Seq. No. 774

Contig ID 419 1.R1040

5'-most EST LIB3049-020-Q1-E1-E6

Method BLASTN
NCBI GI g3021374
BLAST score 361
E value 0.0e+00
Match length 389
% identity 98

NCBI Description Glycine max mRNA for profilin, PRO1

Seq. No. 775

Contig ID 419_2.R1040 5'-most EST leu701145541.h1

Method BLASTN
NCBI GI g556835
BLAST score 309
E value 1.0e-173
Match length 617
% identity 89

NCBI Description P.vulgaris mRNA for profilin

Seq. No. 776

Contig ID 419 3.R1040

5'-most EST LIB3040-004-Q1-E1-E1

Method BLASTX
NCBI GI g2213610
BLAST score 245
E value 1.0e-20



Match length 26 % identity

(AC000103) F21J9.4 [Arabidopsis thaliana] NCBI Description

Seq. No. 777

419 4.R1040 Contig ID

LIB3109-006-Q1-K1-B7 5'-most EST

BLASTN Method g3021372 NCBI GI BLAST score 182 2.0e-97 E value 381 Match length % identity 91

NCBI Description Glycine max mRNA for profilin, PRO2

Seq. No. 778

419 5.R1040 Contig ID trc700567705.h1 5'-most EST

BLASTN Method g3021372 NCBI GI 70 BLAST score 8.0e-31 E value 98 Match length 93 % identity

NCBI Description Glycine max mRNA for profilin, PRO2

779 Seq. No.

419 6.R1040 Contig ID

LIB3107-052-Q1-K1-D1 5'-most EST

BLASTN Method q3021372 NCBI GI 217 BLAST score 1.0e-118 E value 269 Match length

% identity 96

NCBI Description Glycine max mRNA for profilin, PRO2

780 Seq. No.

419 7.R1040 Contig ID hrw701059647.h1 5'-most EST

BLASTN Method g3021374 NCBI GI 229 BLAST score E value 1.0e-126 Match length 349 91 % identity

NCBI Description Glycine max mRNA for profilin, PRO1

781 Seq. No.

419 9.R1040 Contig ID sat701006165.h1 5'-most EST

BLASTN Method g3021374 NCBI GI 142 BLAST score 7.0e-74 E value 229 Match length 91 % identity



NCBI Description Glycine max mRNA for profilin, PRO1

Seq. No. 782
Contig ID 419 10.

Contig ID 419_10.R1040 5'-most EST jsh701070477.h1

Method BLASTN
NCBI GI g3021372
BLAST score 34
E value 2.0e-09
Match length 58
% identity 90

NCBI Description Glycine max mRNA for profilin, PRO2

Seq. No. 783

Contig ID 419_11.R1040 5'-most EST fua701042750.h1

Method BLASTN
NCBI GI g3021374
BLAST score 129
E value 2.0e-66
Match length 137
% identity 99

NCBI Description Glycine max mRNA for profilin, PRO1

Seq. No. 784

Contig ID 419_14.R1040

5'-most EST LIB3072-021-Q1-E1-B8

Method BLASTN
NCBI GI g3021374
BLAST score 83
E value 9.0e-39
Match length 123
% identity 92

% identity 92

NCBI Description Glycine max mRNA for profilin, PRO1

Seq. No. 785

Contig ID 419_21.R1040

5'-most EST LIB3050-021-Q1-K1-C3

Method BLASTN
NCBI GI g3021374
BLAST score 116
E value 1.0e-58
Match length 140
% identity 96

NCBI Description Glycine max mRNA for profilin, PRO1

Seq. No. 786

421 1.R1040 Contig ID 5'-most EST g1321819 FL Method BLASTX g1321820 NCBI GI BLAST score 1854 0.0e+00E value Match length 392 % identity 92

NCBI Description (X96864) phosphoribosylamine--glycine ligase [Glycine max]



Seq. No. 422 1.R1040 Contig ID 5'-most EST g1321821 FL Method BLASTN g1321821 NCBI GI BLAST score 720 0.0e + 00E value 1302 Match length % identity 99 G.max mRNA for glycinamide ribonucleotide transformylase NCBI Description Seq. No. 422 2.R1040 Contig ID 5'-most EST epx701103931.h1 Method BLASTN q1321821 NCBI GI 169 BLAST score 4.0e-90 E value 306 Match length 98 % identity NCBI Description G.max mRNA for glycinamide ribonucleotide transformylase 789 Seq. No. 423 1.R1040 Contig ID 5'-most EST sat701005463.hl BLASTX Method q112837 NCBI GI BLAST score 1630 0.0e+00E value Match length 315 100 % identity NAD(P)H DEPENDENT 6'-DEOXYCHALCONE SYNTHASE NCBI Description >gi_99953_pir__S14222 reductase - soybean >gi_18728_emb_CAA39261_ (X55730) reductase [Glycine max] Seq. No. 790 Contig ID 423 2.R1040 bth700849472.h1 5'-most EST BLASTX Method NCBI GI q112837 BLAST score 1139 1.0e-125 E value Match length 315 66 % identity NAD(P)H DEPENDENT 6'-DEOXYCHALCONE SYNTHASE NCBI Description >gi 99953 pir S14222 reductase - soybean >gi 18728 emb CAA39261 (X55730) reductase [Glycine max] 791 Seq. No. 423 3.R1040 Contig ID uC-gmrominsoy046b12b1 5'-most EST BLASTX Method

g112837 NCBI GI 447 BLAST score 1.0e-50 E value 147 Match length % identity 67

```
NAD(P)H DEPENDENT 6'-DEOXYCHALCONE SYNTHASE
NCBI Description
                  >gi_99953_pir__S14222 reductase - soybean
                  >gi 18728 emb CAA39261 (X55730) reductase [Glycine max]
                  792
Seq. No.
                  423 5.R1040
Contig ID
                  zsg701128036.h1
5'-most EST
Method
                  BLASTN
                  g18727
NCBI GI
                  213
BLAST score
                  1.0e-116
E value
                  320
Match length
% identity
NCBI Description Soybean mRNA for reductase involved in deoxychalcone
                  synthesis (NAD(P)H dependent 6'-deoxychalcone synthase)
                  793
Seq. No.
                  424 1.R1040
Contig ID
5'-most EST
                  g2764803 FL
                  BLASTX
Method
                  g1304227
NCBI GI
                  1849
BLAST score
                  0.0e + 00
E value
Match length
                  341
% identity
                  100
                  (D63781) Epoxide hydrolase [Glycine max]
NCBI Description
                  >gi 2764804_emb_CAA55293_ (X78547) epoxide hydrolase
                  [Glycine max]
                  794
Seq. No.
                  424 2.R1040
Contig ID
```

ssr700559402.h1 5'-most EST

BLASTN Method NCBI GI q1304226 385 BLAST score E value 0.0e + 00Match length 473 % identity

NCBI Description Soybean mRNA for epoxide hydrolase, complete cds

795 Seq. No.

424 3.R1040 Contig ID 5'-most EST crh700850094.h1

BLASTN Method NCBI GI g2764803 BLAST score 209 1.0e-113 E value 805 Match length % identity 90

NCBI Description G.max mRNA for epoxide hydrolase

796 Seq. No.

426 1.R1040 Contig ID

LIB3072-018-Q1-E1-A9 5'-most EST

Method BLASTN g18747 NCBI GI 384 BLAST score



0.0e + 00E value 427 Match length 98 % identity

G.max mRNA for a protein similar to potato tuber protein NCBI Description

p322 homolgous to Bowman-Birk Proteinase Inhibitor

797 Seq. No.

427 1.R1040 Contig ID

LIB3093-005-Q1-K1-E2 5'-most EST

BLASTX Method g3913724 NCBI GI 1851 BLAST score 0.0e + 00E value 373 Match length 96

% identity GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-2 SUBUNIT NCBI Description

(GP-ALPHA-2) >gi_1834453_emb_CAA64834_ (X95582) G protein

[Glycine max]

798 Seq. No.

428 1.R1040 Contig ID

LIB3170-044-Q1-J1-A5 5'-most EST

BLASTN Method g18749 NCBI GI 714 BLAST score 0.0e + 00E value 714 Match length 100 % identity

NCBI Description G.max mRNA for seed maturation polypeptide

799 Seq. No.

428 2.R1040 Contig ID

LIB3051-102-Q1-K1-G3 5'-most EST

BLASTN Method q170007 NCBI GI 836 BLAST score 0.0e + 00E value 836 Match length 100 % identity

NCBI Description Soybean 18 kD late embryogenesis abundant (Lea) protein

mRNA, complete cds

800 Seq. No.

429 1.R1040 Contig ID

LIB3039-021-Q1-E1-A3 5'-most EST

BLASTN Method g1055367 NCBI GI 639 BLAST score E value 0.0e + 00706 Match length % identity 100

Glycine max ribulose-1,5-bisphosphate carboxylase small NCBI Description

subunit mRNA, complete cds

801 Seq. No.

429 2.R1040 Contig ID

LIB3055-005-Q1-N1-G4 5'-most EST



Method BLASTN
NCBI GI g1055367
BLAST score 199
E value 1.0e-108
Match length 371
% identity 94

NCBI Description Glycine max ribulose-1,5-bisphosphate carboxylase small

subunit mRNA, complete cds

Seq. No. 802

Contig ID 429_4.R1040 5'-most EST wrg700786845.h2

Method BLASTN
NCBI GI g1055367
BLAST score 155
E value 1.0e-81
Match length 175
% identity 98

NCBI Description Glycine max ribulose-1,5-bisphosphate carboxylase small

subunit mRNA, complete cds

Seq. No. 803

Contig ID 429_9.R1040

5'-most EST LIB3074-039-Q1-K1-F4

Method BLASTN
NCBI GI g1055367
BLAST score 222
E value 1.0e-121
Match length 521
% identity 95

NCBI Description Glycine max ribulose-1,5-bisphosphate carboxylase small

subunit mRNA, complete cds

Seq. No. 804

Contig ID 429_10.R1040

5'-most EST LIB3106-110-Q1-K1-A9

Method BLASTN
NCBI GI g1055367
BLAST score 224
E value 1.0e-123
Match length 396
% identity 90

NCBI Description Glycine max ribulose-1,5-bisphosphate carboxylase small

subunit mRNA, complete cds

Seq. No. 805

Contig ID 431_1.R1040

5'-most EST LIB3107-037-Q1-K1-C4

Method BLASTN
NCBI GI g18761
BLAST score 988
E value 0.0e+00
Match length 988
% identity 100

NCBI Description Soybean stem mRNA for 31 kD glycoprotein

Seq. No. 806



```
431 2.R1040
Contig ID
                  g169897_FL
5'-most EST
Method
                  BLASTN
                  q169897
NCBI GI
                  1074
BLAST score
                  0.0e+00
E value
                  1074
Match length
                  100
% identity
NCBI Description G.max 28 kDa protein, complete cds
                  807
Seq. No.
                  431 4.R1040
Contig ID
5'-most EST
                  taw700658005.h1
Method
                  BLASTN
                  g1079735
NCBI GI
                  111
BLAST score
                  1.0e-55
E value
                  115
Match length
                  99
% identity
                  Glycine soja ribulose 1,5-bisphosphate carboxylase small
NCBI Description
                  subunit precursor (rbcS) gene, nuclear gene encoding
                  chloroplast protein, complete cds
                   808
Seq. No.
                   431 5.R1040
Contig ID
                  LIB3106-051-Q1-K1-B5
5'-most EST
                  BLASTN
Method
                   q18761
NCBI GI
BLAST score
                   175
                   2.0e-93
E value
                   425
Match length
                   85
% identity
NCBI Description Soybean stem mRNA for 31 kD glycoprotein
                   809
Seq. No.
                   431 6.R1040
Contig ID
5'-most EST
                   LIB3039-029-Q1-E1-H3
                   BLASTN
Method
NCBI GI
                   q169974
BLAST score
                   170
E value
                   1.0e-90
Match length
                   323
                   98
% identity
NCBI Description Glycine max vspA gene, complete cds
```

810 Seq. No.

431 7.R1040 Contig ID

5'-most EST LIB3106-091-Q1-K1-D9

Method BLASTN NCBI GI g170089 394 BLAST score 0.0e + 00E value 518 Match length 98 % identity

NCBI Description G.max vegetative storage protien mRNA (VSP27), complete cds

811 Seq. No.



431 8.R1040 Contig ID 5'-most EST $q50\overline{5}8414$ BLASTN Method q169897 NCBI GI BLAST score 218 1.0e-119 E value Match length 683 % identity 95 NCBI Description G.max 28 kDa protein, complete cds 812 Seq. No. Contig ID 431 9.R1040 5'-most EST LIB3106-064-Q1-K1-D8 Method BLASTN g170091 NCBI GI 283 BLAST score 1.0e-158 E value Match length 299 % identity 99 NCBI Description Glycine max vegetative storage protein (vspB) gene, complete cds Seq. No. 813 Contig ID 431 11.R1040 5'-most EST LIB3039-035-Q1-E1-F7 Method BLASTN NCBI GI q169897 BLAST score 266 E value 1.0e-148 Match length 568 % identity NCBI Description G.max 28 kDa protein, complete cds Seq. No. 814 432 1.R1040 Contig ID 5'-most EST g256428 FL Method BLASTN NCBI GI g256428 BLAST score 1212 0.0e+00E value Match length 1264 % identity 99

KTi3 (KTi3+)=Kunitz trypsin inhibitor [soybeans, Dare and NCBI Description Forrest, mRNA, 1259 nt]

Seq. No. 815

432_5.R1040 Contig ID

5'-most EST LIB3072-020-Q1-E1-G8

Method BLASTN NCBI GI q256428 BLAST score 324 0.0e+00E value Match length 406 96 % identity

NCBI Description KTi3 (KTi3+)=Kunitz trypsin inhibitor [soybeans, Dare and

Forrest, mRNA, 1259 nt]



Seq. No.

434 1.R1040 Contig ID

5'-most EST LIB3051-050-Q1-K1-D12

817 Seq. No.

434 2.R1040 Contig ID

LIB3051-049-Q1-K1-F6 5'-most EST

BLASTN Method g312988 NCBI GI 226 BLAST score 1.0e-124 E value 414 Match length % identity 89

NCBI Description G.max mRNA for beta-tubulin, partial cds

818 Seq. No.

434 3.R1040 Contig ID

LIB3039-007-Q1-E1-C3 5'-most EST

BLASTN Method g1399449 NCBI GI BLAST score 147 7.0e-77 E value 367 Match length 86 % identity

NCBI Description Lupinus albus beta-tubulin 2 (TubB2) gene, complete cds

819 Seq. No.

434 4.R1040 Contig ID 5'-most EST ncj700983044.h1

Method BLASTN g1403142 NCBI GI 454 BLAST score 0.0e + 00E value Match length 1207 % identity

NCBI Description C.arietinum mRNA for beta-tubulin

820 Seq. No.

Contig ID 434 5.R1040

LIB3087-010-Q1-K1-E10 5'-most EST

Method BLASTX g3915083 NCBI GI BLAST score 1635 E value 0.0e + 00423 Match length 71 % identity

TUBULIN BETA CHAIN >gi 1403143 emb CAA67056 (X98406) NCBI Description

beta-tubulin [Cicer arietinum]

821 Seq. No.

434 6.R1040 Contig ID

LIB3049-011-Q1-E1-A4 5'-most EST

BLASTX Method g1351202 NCBI GI BLAST score 749 2.0e-79 E value 237 Match length



% identity 52
NCBI Description TUBULIN BETA CHAIN >gi_312989_emb_CAA42777_ (X60216)
beta-tubulin [Glycine max]

Seq. No. 822

434 7.R1040 Contig ID q42895995'-most EST BLASTN Method g312988 NCBI GI 1134 BLAST score 0.0e+00E value 1234 Match length 98 % identity

NCBI Description G.max mRNA for beta-tubulin, partial cds

Seq. No. 823

Contig ID 434 8.R1040

5'-most EST uC-gmflminsoy118b09b1

Method BLASTN
NCBI GI g170061
BLAST score 327
E value 0.0e+00
Match length 495
% identity 96

NCBI Description Soybean (G.max L.) beta-tubulin (S-beta-2) gene, complete

cds

Seq. No. 824

Contig ID 434_9.R1040

5'-most EST jC-gmf102220136d01d1

Method BLASTN
NCBI GI g4098318
BLAST score 123
E value 1.0e-62
Match length 375
% identity 84

NCBI Description Triticum aestivum beta-tubulin 1 (tubb1) mRNA, complete cds

Seq. No. 825

435 1.R1040 Contig ID g392820 FL 5'-most EST BLASTN Method q392820 NCBI GI 1045 BLAST score E value 0.0e + 001045 Match length 100 % identity

NCBI Description Glycine max Williams cellulase mRNA, partial cds

Seq. No. 826

Contig ID 437 1.R1040

5'-most EST LIB3109-001-Q1-K1-C11

Method BLASTN
NCBI GI g555615
BLAST score 1040
E value 0.0e+00
Match length 1040



% identity 100

NCBI Description Glycine max Mandarin cytokinin induced message (cim1) mRNA,

complete cds

Seq. No. 827

Contig ID 438_1.R1040 5'-most EST jsh701065288.h1

Method BLASTX
NCBI GI g1170214
BLAST score 1876
E value 0.0e+00
Match length 412
% identity 90

NCBI Description DELTA-AMINOLEVULINIC ACID DEHYDRATASE PRECURSOR

(PORPHOBILINOGEN SYNTHASE) (ALADH) >gi_468000 (U04525) delta-aminolevulinic acid dehydratase [Glycine max]

Seq. No. 828

Contig ID 441 1.R1040

5'-most EST LIB3139-110-P1-N1-F10

Method BLASTN
NCBI GI 9472849
BLAST score 479
E value 0.0e+00
Match length 551
% identity 97

NCBI Description Glycine max Essex desiccation protectant protein Lea14

homolog mRNA, complete cds

Seq. No. 829

Contig ID 441_2.R1040 5'-most EST dpv701101866.h1

Method BLASTN
NCBI GI g472849
BLAST score 150
E value 1.0e-78
Match length 198
% identity 94

NCBI Description Glycine max Essex desiccation protectant protein Lea14

homolog mRNA, complete cds

Seq. No. 830

Contig ID 442_1.R1040 5'-most EST gsv701049208.h1

Method BLASTX
NCBI GI g2642238
BLAST score 2675
E value 0.0e+00
Match length 668
% identity 91

NCBI Description (AF031241) endoplasmic reticulum HSC70-cognate binding

protein precursor [Glycine max]

Seq. No. 831

Contig ID 442_2.R1040 5'-most EST g475601_FL Method BLASTX



```
q475602
NCBI GI
BLAST score
                  3011
E value
                  0.0e+00
Match length
                  642
% identity
                  93
                  (U08384) BiP isoform A [Glycine max]
NCBI Description
                  832
Seq. No.
                   442 3.R1040
Contig ID
5'-most EST
                   k11701206325.h1
Method
                  BLASTX
                   g2924518
NCBI GI
BLAST score
                   316
E value
                   2.0e-28
Match length
                   189
% identity
                   37
                  (AL022023) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   833
                   442 4.R1040
Contig ID
5'-most EST
                   LIB3030-012-Q1-B1-E11
Seq. No.
                   834
                   442 7.R1040
Contig ID
                   ssr700558430.hl
5'-most EST
Method
                   BLASTN
                   g475597
NCBI GI
BLAST score
                   161
                   2.0e-85
E value
                   254
Match length
                   97
% identity
                  Glycine max Century 84 BiP isoform C mRNA, partial cds
NCBI Description
                   835
Seq. No.
                                                187
                   444 1.R1040
Contig ID
                   LIB3051-020-Q1-E1-G10
5'-most EST
Method
                   BLASTN
                   q475603
NCBI GI
                   769
BLAST score
                   0.0e + 00
E value
                   919
Match length
                   99
% identity
NCBI Description Glycine max Century 84 beta-1,3-glucanase mRNA, partial cds
                   836
Seq. No.
                   444 2.R1040
Contig ID
                   LIB3051-027-Q1-K1-G3
5'-most EST
                   BLASTN
Method
NCBI GI
                   q475603
BLAST score
                   346
E value
                   0.0e+00
Match length
                   615
```

837 Seq. No.

NCBI Description

% identity

445 1.R1040 Contig ID

Glycine max Century 84 beta-1,3-glucanase mRNA, partial cds

NCBI Description

cds



```
fua701037217.h1
5'-most EST
                  BLASTX
Method
                  g550452
NCBI GI
BLAST score
                  3467
                  0.0e + 00
E value
                  674
Match length
                  100
% identity
                   (U08469) 3-methylcrotonyl-CoA carboxylase, biotin-carrier
NCBI Description
                  domain [Glycine max]
                   838
Seq. No.
                   448 1.R1040
Contig ID
                  LIB3052-013-Q1-N1-A10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g267076
BLAST score
                   1628
                   0.0e+00
E value
                   334
Match length
                   93
% identity
                  TUBULIN BETA-3 CHAIN >gi_81953_pir__S20870 tubulin beta-3
NCBI Description
                   chain - garden pea (fragment) >gi_388255_emb_CAA38615_
                   (X54846) beta-tubulin 3 [Pisum sativum]
                   839
Seq. No.
                   448 2.R1040
Contig ID
                   ncj700978473.h1
5'-most EST
                   BLASTN
Method
                   g1399449
NCBI GI
                   370
BLAST score
                   0.0e + 00
E value
                   789
Match length
                   90
% identity
                  Lupinus albus beta-tubulin 2 (TubB2) gene, complete cds
NCBI Description
                   840
Seq. No.
Contig ID
                   448 3.R1040
5'-most EST
                   LIB3049-023-Q1-E1-F11
Method
                   BLASTX
                   q2983997
NCBI GI
BLAST score
                   348
                   1.0e-34
E value
                   159
Match length
% identity
                   (AE000749) hypothetical protein [Aquifex aeolicus]
NCBI Description
Seq. No.
                   448 4.R1040
Contig ID
5'-most EST
                   pmv700892602.hl
Method
                   BLASTN
NCBI GI
                   q170061
BLAST score
                   545
                   0.0e + 00
E value
                   932
Match length
                   94
% identity
```

Soybean (G.max L.) beta-tubulin (S-beta-2) gene, complete



```
Seq. No.
                  448 5.R1040
Contig ID
                  LIB\overline{3}040-037-Q1-E1-B9
5'-most EST
Method
                  BLASTX
                  q2983997
NCBI GI
                  153
BLAST score
                  5.0e-10
E value
                  53
Match length
% identity
                  51
                  (AE000749) hypothetical protein [Aquifex aeolicus]
NCBI Description
                  843
Seq. No.
                   448 6.R1040
Contig ID
                  LIB3170-026-Q1-K1-H6
5'-most EST
Method
                  BLASTN
                   g515691
NCBI GI
BLAST score
                   46
                   9.0e-17
E value
                   70
Match length
                   91
% identity
                  Glycine max Forrest beta-tubulin (tubB3) mRNA, partial cds
NCBI Description
                   844
Seq. No.
                   448 7.R1040
Contig ID
5'-most EST
                   LIB3170-065-Q1-J1-E6
                   BLASTN
Method
                   g1399449
NCBI GI
                   78
BLAST score
                   2.0e-35
E value
Match length
                   141
                   89
% identity
NCBI Description Lupinus albus beta-tubulin 2 (TubB2) gene, complete cds
                   845
Seq. No.
                   448 9.R1040
Contig ID
5'-most EST
                   ncj700979970.hl
                   BLASTN
Method
                   q515691
NCBI GI
BLAST score
                   218
                   1.0e-119
E value
                   302
Match length
                   93
% identity
NCBI Description Glycine max Forrest beta-tubulin (tubB3) mRNA, partial cds
Seq. No.
                   846
                   448 10.R1040
Contig ID
                   sat701013054.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g170061
                   235
BLAST score
E value
                   1.0e-129
                   416
Match length
                   97
 % identity
                   Soybean (G.max L.) beta-tubulin (S-beta-2) gene, complete
NCBI Description
```

Seq. No. 847

cds



```
Contig ID
                   448 11.R1040
5'-most EST
                  uC-gmflminsoy019d02b1
                   848
Seq. No.
                   448 13.R1040
Contig ID
                   zhf700953739.h1
5'-most EST
                  BLASTN
Method
                   q515691
NCBI GI
BLAST score
                  294
                   1.0e-164
E value
                   464
Match length
                   95
% identity
NCBI Description Glycine max Forrest beta-tubulin (tubB3) mRNA, partial cds
                   849
Seq. No.
                   449 1.R1040
Contig ID
5'-most EST
                   pcp700990393.hl
Method
                  BLASTX
NCBI GI
                   q530088
BLAST score
                   2101
                   0.0e + 00
E value
                   389
Match length
                   100
% identity
                  (U12735) aminoalcoholphosphotransferase [Glycine max]
NCBI Description
Seq. No.
                   850
                   449 2.R1040
Contig ID
5'-most EST
                   LIB3093-040-Q1-K1-C9
Method
                   BLASTN
                   q530086
NCBI GI
                   231
BLAST score
                   1.0e-127
E value
Match length
                   283
% identity
                   95
NCBI Description
                  Glycine max aminoalcoholphosphotransferase (AAPT1) mRNA,
                   complete cds
Seq. No.
                   851
Contig ID
                   450 1.R1040
5'-most EST
                   g927574 FL
Method
                   BLASTX
NCBI GI
                   g927575
                   2221
BLAST score
E value
                   0.0e + 00
                   422
Match length
% identity
                   98
NCBI Description
                  (U12926) alpha galactosidase [Glycine max]
                   852
Seq. No.
                   451 1.R1040
Contig ID
                   LIB3139-012-P1-N1-F8
```

5'-most EST

Method BLASTX g2370312 NCBI GI BLAST score 1779 E value 0.0e + 00Match length 420 % identity 81



NCBI Description (AJ000995) DnaJ-like protein [Medicago sativa] >gi_3202020 (AF069507) DnaJ-like protein MsJ1 [Medicago sativa]

Seq. No. 853

Contig ID 451 2.R1040

5'-most EST jC-gmle01810083a09d1

Method BLASTN
NCBI GI g532700
BLAST score 209
E value 1.0e-114
Match length 261
% identity 95

NCBI Description Glycine max clone GMFP1 isoprenylated protein mRNA, partial

cds

Seq. No. 854

Contig ID 451 3.R1040

5'-most EST uC-gmropic071d07b1

Method BLASTN
NCBI GI g2370311
BLAST score 33
E value 8.0e-09
Match length 104
% identity 89

NCBI Description Medicago sativa mRNA for DnaJ-like protein

Seq. No. 855

Contig ID 451_5.R1040

5'-most EST jC-gmro02910030f08d1

Method BLASTN
NCBI GI g2370311
BLAST score 131
E value 2.0e-67
Match length 263
% identity 87

NCBI Description Medicago sativa mRNA for DnaJ-like protein

Seq. No. 856

Contig ID 452 1.R1040

5'-most EST LIB3106-100-Q1-K1-A11

Method BLASTN
NCBI GI g532702
BLAST score 372
E value 0.0e+00
Match length 428
% identity 92

NCBI Description Glycine max clone GMFP2 isoprenylated protein mRNA, partial

cds

Seq. No. 857

Contig ID 452 2.R1040

5'-most EST LIB3092-062-Q1-K1-H7

Method BLASTX
NCBI GI g123620
BLAST score 2996
E value 0.0e+00
Match length 618

4 (45 · 15



% identity NCBI Description

HEAT SHOCK COGNATE 70 KD PROTEIN 2 >gi 100224 pir S14950

heat shock cognate protein 70 - tomato

>gi_19258_emb_CAA37971_ (X54030) heat shock protein cognate

70 [Lycopersicon esculentum]

858 Seq. No.

452 3.R1040 Contig ID 5'-most EST bth700848091.h1

BLASTN Method g532704 NCBI GI BLAST score 465 0.0e + 00E value Match length 473 % identity 78

Glycine max clone GMFP3 isoprenylated protein mRNA, partial NCBI Description

cds

859 Seq. No.

452 4.R1040 Contig ID sat701012856.h1 5'-most EST

860 Seq. No.

452 5.R1040 Contig ID 5'-most EST jex700907339.hl

Seq. No. 861

452 8.R1040 Contig ID

5'-most EST LIB3039-040-Q1-E1-H12

Method BLASTX g1565257 NCBI GI 194 BLAST score 2.0e-14 E value Match length 170

% identity

(X57108) cerebroside sulfate activator [Homo sapiens] NCBI Description

862 Seq. No.

452 10.R1040 Contig ID

5'-most EST uC-gmflminsoy064b08b1

Method BLASTX NCBI GI g3043415 BLAST score 230 E value 3.0e-31 Match length 92 % identity 80

(Y17053) At-hsc70-3 [Arabidopsis thaliana] NCBI Description

863 Seq. No.

452 14.R1040 Contig ID 5'-most EST q5510286 BLASTN Method g1430886 NCBI GI 150 BLAST score 2.0e-78 E value 378 Match length % identity 85





NCBI Description M.sativa mRNA for 70 kD heat shock protein

Seq. No. 864

Contig ID 454_1.R1040 5'-most EST pcp700989850.h1

Method BLASTN
NCBI GI g532706
BLAST score 200
E value 1.0e-108
Match length 249
% identity 99

% identity 99
NCBI Description Glycine max clone GMFP6 isoprenylated protein mRNA, partial

cds

Seq. No. 865

Contig ID 455_1.R1040

5'-most EST uC-gmrominsoy298d08b1

Method BLASTX
NCBI GI g4389417
BLAST score 4507
E value 0.0e+00
Match length 890
% identity 94

NCBI Description (AF055369) nitrate reductase [Glycine max]

Seq. No. 866

Contig ID 455_2.R1040

5'-most EST kmv700739087.h1

Method BLASTX
NCBI GI g4337177
BLAST score 1687
E value 0.0e+00
Match length 376
% identity 84

NCBI Description (AC006416) Identical to gb_Y10557 g5bf gene from

Arabidopsis thaliana. ESTs gb_R30578, gb_R90475,

gb_T22384, gb_T22425, gb_N6493 $\overline{4}$ and gb_T4 $\overline{6}$ 767 come from

this gene. [Arabidopsis thaliana]

Seq. No. 867

Contig ID 455_3.R1040 5'-most EST fua701042206.h1

Method BLASTX
NCBI GI g1709267
BLAST score 4552
E value 0.0e+00
Match length 886
% identity 95

NCBI Description INDUCIBLE NITRATE REDUCTASE 1 (NR) >gi_1262166 (L23854)

nitrate reductase [Glycine max]

Seq. No. 868

Contig ID 455_4.R1040

5'-most EST jC-gmro02910013d05d1

Method BLASTX
NCBI GI g1906830
BLAST score 1531



E value 1.0e-171 Match length 324

% identity 90

NCBI Description (Y11829) heat shock protein [Arabidopsis thaliana]

Seq. No. 869

Contig ID 455_5.R1040

5'-most EST jC-gmle01810080a04a1

Method BLASTX
NCBI GI g1076758
BLAST score 1114
E value 1.0e-122
Match length 272
% identity 77

NCBI Description heat-shock protein precursor - rye >gi_2130093_pir__S65776

heat-shock protein, 82K, precursor - rye

>gi 556673_emb_CAA82945_ (Z30243) heat-shock protein

[Secale cereale]

Seq. No. 870

Contig ID 455_7.R1040

5'-most EST jC-gmle01810082h03d1

Method BLASTN
NCBI GI g1262165
BLAST score 130
E value 5.0e-67
Match length 146
% identity 97

NCBI Description Soybean nitrate reductase (INR1) mRNA, complete cds

Seq. No. 871

Contig ID 457 1.R1040

5'-most EST LIB3051-081-Q1-K1-C1

Method BLASTX
NCBI GI g1705678
BLAST score 4065
E value 0.0e+00
Match length 807
% identity 98

NCBI Description CELL DIVISION CYCLE PROTEIN 48 HOMOLOG (VALOSIN CONTAINING

PROTEIN HOMOLOG) (VCP) >gi_862480 (U20213) valosin-containing protein [Glycine max]

Seq. No. 872

Contig ID 457_2.R1040

5'-most EST fC-gmro700745012a1

Method BLASTX
NCBI GI 94432856
BLAST score 539
E value 6.0e-55
Match length 214
% identity 49

NCBI Description (AC006300) putative 2A6 protein [Arabidopsis thaliana]

Seq. No. 873

Contig ID 457_3.R1040 5'-most EST epx701107517.h1



```
BLASTN
Method
NCBI GI
                  q862479
BLAST score
                  145
                  2.0e-75
E value
                  221
Match length
% identity
NCBI Description Glycine max valosin-containing protein mRNA, complete cds
                  874
Seq. No.
                  457 4.R1040
Contig ID
5'-most EST
                  crh700851542.h1
                  875
Seq. No.
                  457 5.R1040
Contig ID
                  LIB3170-060-Q1-J1-F8
5'-most EST
                  BLASTN
Method
                   g862479
NCBI GI
                   252
BLAST score
                   1.0e-139
E value
                   415
Match length
                   90
% identity
NCBI Description Glycine max valosin-containing protein mRNA, complete cds
Seq. No.
                   876
                   457 6.R1040
Contig ID
                   LIB3170-060-Q1-K1-F8
5'-most EST
                   BLASTN
Method
                   q862479
NCBI GI
BLAST score
                   85
                   5.0e-40
E value
                   133
Match length
% identity
NCBI Description Glycine max valosin-containing protein mRNA, complete cds
Seq. No.
                   457 9.R1040
Contig ID
                   g4302765
5'-most EST
                   BLASTN
Method
NCBI GI
                   q862479
BLAST score
                   415
                   0.0e + 00
E value
                   597
Match length
                   95
% identity
NCBI Description Glycine max valosin-containing protein mRNA, complete cds
                   878
Seq. No.
                   458 1.R1040
Contig ID
                   ncj700986139.hl
5'-most EST
                   BLASTX
Method
                   g3334138
NCBI GI
                   2332
BLAST score
                   0.0e+00
E value
```

[Glycine max]

510

87

Match length

NCBI Description

% identity

CALNEXIN HOMOLOG PRECURSOR >gi_669003 (U20502) calnexin



```
Seq. No.
                  458 2.R1040
Contig ID
                  LIB3170-056-Q1-K1-G8
5'-most EST
Method
                  BLASTN
                  g669002
NCBI GI
                  183
BLAST score
                  3.0e-98
E value
                  412
Match length
                  86
% identity
NCBI Description Glycine max calnexin mRNA, complete cds
                  880
Seq. No.
                  458 3.R1040
Contig ID
                  leu701147326.hl
5'-most EST
Method
                  BLASTX
                  g3334138
NCBI GI
BLAST score
                  600
                  6.0e-88
E value
                  243
Match length
                  69
% identity
                  CALNEXIN HOMOLOG PRECURSOR >gi_669003 (U20502) calnexin
NCBI Description
                   [Glycine max]
                   881
Seq. No.
                   458 4.R1040
Contig ID
5'-most EST
                   g4301841
                   BLASTN
Method
                   g669002
NCBI GI
BLAST score
                   290
                   1.0e-162
E value
                   642
Match length
                   92
% identity
NCBI Description Glycine max calnexin mRNA, complete cds
                   882
Seq. No.
                   458 7.R1040
Contig ID
5'-most EST
                   ncj700975503.hl
Method
                   BLASTN
                   q669002
NCBI GI
BLAST score
                   161
E value
                   3.0e-85
                   303
Match length
                   100
% identity
NCBI Description Glycine max calnexin mRNA, complete cds
                   883
Seq. No.
                   458 9.R1040
Contig ID
                   LIB3051-082-Q1-K1-F9
5'-most EST
Method
                   BLASTN
```

Method BLASIN
NCBI GI g669002
BLAST score 69
E value 2.0e-30
Match length 169
% identity 85

NCBI Description Glycine max calnexin mRNA, complete cds

Seq. No. 884



459 1.R1040 Contig ID 5'-most EST LIB3138-014-Q1-N2-C2 BLASTN Method q710431 NCBI GI 918 BLAST score E value 0.0e + 00Match length 948 99 % identity Glycine max low molecular weight heat shock protein Hsp23.9 NCBI Description (Gmhsp23.9) mRNA, complete cds 885 Seq. No. Contig ID 459 2.R1040 5'-most EST LIB3051-016-Q1-E1-C10 BLASTN Method q710431 NCBI GI 248 BLAST score E value 1.0e-137 252 Match length 100 % identity Glycine max low molecular weight heat shock protein Hsp23.9 NCBI Description (Gmhsp23.9) mRNA, complete cds Seq. No. 886 460 1.R1040 Contig ID LIB3107-065-Q1-K1-D5 5'-most EST Method BLASTN q710433 NCBI GI BLAST score 361 0.0e+00E value Match length 707 99 % identity Glycine max low molecular weight heat shock protein Hsp22.3 NCBI Description (Gmhsp22.3) mRNA, complete cds 887 Seq. No. 461 1.R1040 Contig ID hrw701058941.h1 5'-most EST Method BLASTN g710435 NCBI GI BLAST score 723 0.0e+00E value 727 Match length % identity 100 Glycine max low molecular weight heat shock protein Hsp22.5 NCBI Description (Gmhsp22.5) mRNA, complete cds 888 Seq. No. 462_1.R1040 Contig ID uC-gmrominsoy194b09b1 5'-most EST Method BLASTX NCBI GI g2129826 BLAST score 3024

0.0e+00E value Match length 610 98 % identity

NCBI Description dynamin-like protein phragmoplastin 5 - soybean >gi_1218004

Method NCBI GI

E value

BLAST score

Match length



(U36430) SDL5A [Glycine max]

```
Seq. No.
                   889
                   462_2.R1040
Contig ID
                   g1217993 FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2129825
BLAST score
                   3029
                   0.0e + 00
E value
                   610
Match length
                   98
% identity
                   dynamin-like protein phragmoplastin 12 - soybean
NCBI Description
                   >gi 1217994 (U25547) SDL [Glycine max]
                   890
Seq. No.
                   462 3.R1040
Contig ID
5'-most EST
                   jC-\overline{g}mro02910024d01d1
Method
                   BLASTN
NCBI GI
                   q1217993
BLAST score
                   118
                   1.0e-59
E value
                   258
Match length
                   86
% identity
                  Glycine max dynamin-like protein SDL12A mRNA, complete cds
NCBI Description
                   891
Seq. No.
                   463 1.R1040
Contig ID
                   sat701012537.h1
5'-most EST
Method
                   BLASTN
                   q886099
NCBI GI
                   983
BLAST score
                   0.0e+00
E value
                   1196
Match length
% identity
                   Glycine max putative water channel protein (Pip1) mRNA,
NCBI Description
                   complete cds
                   892
Seq. No.
Contig ID
                   463 2.R1040
                   leu701151271.h1
5'-most EST
Method
                   BLASTN
                   g886099
NCBI GI
BLAST score
                   220
E value
                   1.0e-120
Match length
                   582
% identity
                   84
                   Glycine max putative water channel protein (Pip1) mRNA,
NCBI Description
                   complete cds
                   893
Seq. No.
                   463 3.R1040
Contig ID
                   LIB3039-019-Q1-E1-A2
5'-most EST
                   BLASTN
```

g886099

204 1.0e-111

312



Seq. No. 894

Contig ID 463 4.R1040

5'-most EST LIB3028-013-Q1-B1-E3

Method BLASTX
NCBI GI g3158476
BLAST score 829
E value 5.0e-89
Match length 176
% identity 91

NCBI Description (AF067185) aquaporin 2 [Samanea saman]

Seq. No. 895

Contig ID 463 5.R1040

5'-most EST LIB3109-035-Q1-K1-C7

Method BLASTN
NCBI GI g886099
BLAST score 339
E value 0.0e+00
Match length 426
% identity 96

NCBI Description Glycine max putative water channel protein (Pip1) mRNA,

complete cds

Seq. No. 896

Contig ID 463_6.R1040 5'-most EST hyd700730567.h1

Method BLASTN
NCBI GI g886099
BLAST score 157
E value 1.0e-82
Match length 369
% identity 92

NCBI Description Glycine max putative water channel protein (Pip1) mRNA,

complete cds

Seq. No. 897

Contig ID 463 7.R1040

5'-most EST LIB3049-044-Q1-E1-B10

Method BLASTN
NCBI GI g886099
BLAST score 278
E value 1.0e-155
Match length 477
% identity 96

NCBI Description Glycine max putative water channel protein (Pip1) mRNA,

complete cds

Seq. No. 898

Contig ID 463_8.R1040 5'-most EST sat701004493.h1

Method BLASTN
NCBI GI g886099
BLAST score 144



E value 3.0e-75
Match length 230
% identity 93

NCBI Description Glycine max putative water channel protein (Pip1) mRNA,

complete cds

Seq. No. 899

Contig ID 463 9.R1040

5'-most EST LIB3109-001-Q1-K1-G12

Method BLASTN
NCBI GI g886099
BLAST score 77
E value 3.0e-35
Match length 104
% identity 94

NCBI Description Glycine max putative water channel protein (Pipl) mRNA,

complete cds

Seq. No. 900

Contig ID 464 1.R1040

5'-most EST fC-gmst700605267f5

Method BLASTN
NCBI GI g1234899
BLAST score 605
E value 0.0e+00
Match length 1049
% identity 91

NCBI Description G.max mRNA for homeobox-leucine zipper protein

Seq. No. 901

464 2.R1040 Contig ID 5'-most EST g5666790 Method BLASTN g1234899 NCBI GI BLAST score 313 1.0e-175 E value 753 Match length 94 % identity

NCBI Description G.max mRNA for homeobox-leucine zipper protein

Seq. No. 902

Contig ID 464_3.R1040 5'-most EST jex700909715.h1

Method BLASTN
NCBI GI g1234899
BLAST score 549
E value 0.0e+00
Match length 553
% identity 100

NCBI Description G.max mRNA for homeobox-leucine zipper protein

Seq. No. 903

Contig ID 465_1.R1040

5'-most EST LIB3051-069-Q1-K1-A2

Method BLASTN NCBI GI g945086 BLAST score 944



E value 0.0e+00 Match length 1054 % identity 97

NCBI Description Glycine max transcription factor TFIIB mRNA, complete cds

Seq. No. 904

Contig ID 465 2.R1040

5'-most EST LIB3170-030-Q1-K1-D2

Method BLASTN
NCBI GI g945086
BLAST score 246
E value 1.0e-136
Match length 272
% identity 98

NCBI Description Glycine max transcription factor TFIIB mRNA, complete cds

Seq. No. 905

Contig ID 466_1.R1040

5'-most EST uC-gmflminsoy069d12b1

Method BLASTN
NCBI GI g1208536
BLAST score 941
E value 0.0e+00
Match length 949
% identity 100

NCBI Description Glycine max guanine nucleotide regulatory protein (rab2)

mRNA, complete cds

Seq. No. 906

Contig ID 466_2.R1040 5'-most EST zzp700836092.h1

Method BLASTN
NCBI GI g1208536
BLAST score 283
E value 1.0e-158
Match length 319
% identity 97

NCBI Description Glycine max guanine nucleotide regulatory protein (rab2)

mRNA, complete cds

Seq. No. 907

467 1.R1040 Contig ID g1161925 FL 5'-most EST Method BLASTX NCBI GI g1161926 BLAST score 3255 E value 0.0e + 00Match length 689 % identity 93

NCBI Description (U34392) alpha-carboxyltransferase aCT-1 precursor [Glycine

max]

Seq. No. 908

Contig ID 467_2.R1040 5'-most EST g1143321_FL Method BLASTX NCBI GI g1143322



BLAST score E value 0.0e + 00Match length 670 95 % identity

(U40979) alfa-carboxyltransferase precursor [Glycine max] NCBI Description

Seq. No.

468 1.R1040 Contig ID 5'-most EST epx701106015.h1

BLASTN Method NCBI GI q506628 BLAST score 650 E value 0.0e + 00901 Match length 96 % identity

Glycine max cv. Dare photosystem II type I chlorophyll NCBI Description

a/b-binding protein (lhcb1*7) gene, complete cds

Seq. No.

468 2.R1040 Contig ID

5'-most EST LIB3039-038-Q1-E1-A12

BLASTN Method NCBI GI q1053215 BLAST score 579 0.0e+00 E value Match length 990 92 % identity

NCBI Description Glycine max chlorophyll a/b-binding protein (cab3) mRNA,

nuclear gene encoding chloroplast protein, complete cds

911 Seq. No.

468 3.R1040 Contig ID

5'-most EST LIB3039-027-Q1-E1-C12

Method BLASTN NCBI GI g18551 BLAST score 522 0.0e + 00E value 844 Match length % identity 95

Soybean Cab3 gene for PSII LHCII chlorophyll a/b binding NCBI Description

protein

Seq. No. 912

Contig ID 471 1.R1040 5'-most EST g1399302 FL Method BLASTX NCBI GI q1399303 2911 BLAST score 0.0e+00E value 549 Match length

% identity NCBI Description (U41473) phosphoinositide-specific phospholipase C P12

[Glycine max]

Seq. No. 913

472 1.R1040 Contig ID 5'-most EST leu701146691.h1

100



99

% identity

NCBI Description

```
Method
                   BLASTX
NCBI GI
                   q1399305
BLAST score
                   3091
                   0.0e + 00
E value
                   600
Match length
                   97
% identity
                   (U41474) phosphoinositide-specific phospholipase C P13
NCBI Description
                   [Glycine max]
                   914
Seq. No.
                   472 2.R1040
Contig ID
                   kl1701212082.hl
5'-most EST
Method
                   BLASTX
                   g1399307
NCBI GI
BLAST score
                   2378
                   0.0e+00
E value
                   627
Match length
                   80
% identity
                   (U41475) phosphoinositide-specific phospholipase C P25
NCBI Description
                   [Glycine max]
                   915
Seq. No.
                   472 3.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy197g12b1
Method
                   BLASTX
NCBI GI
                   q945039
BLAST score
                   3122
                   0.0e+00
E value
Match length
                   600
% identity
                   98
                   (U25027) phosphatidylinositol-specific phospholipase C
NCBI Description
                   [Glycine max]
Seq. No.
                   916
                   474 1.R1040
Contig ID
5'-most EST
                   g5753651
Method
                   BLASTN
                   g1125103
NCBI GI
BLAST score
                   996
                   0.0e+00
E value
Match length
                   1015
% identity
                   100
                   Glycine max seed coat peroxidase isozyme (SPOD4.1) mRNA,
NCBI Description
                   partial cds
Seq. No.
                   917
                   475 1.R1040
Contig ID
                   g3258669
5'-most EST
                   BLASTX
Method
                   g1399380
NCBI GI
BLAST score
                   1933
                   0.0e+00
E value
Match length
                   367
```

methyltransferase [Glycine max]

(U43683) S-adenosyl-L-methionine:delta24-sterol-C-



```
Seq. No. 918
```

Contig ID 475 2.R1040

5'-most EST LIB3049-011-Q1-E1-D2

Method BLASTX
NCBI GI g4185509
BLAST score 614
E value 1.0e-63
Match length 139
% identity 82

NCBI Description (AF102821) actin depolymerizing factor 3 [Arabidopsis

thaliana]

Seq. No. 919

Contig ID 475_3.R1040

5'-most EST LIB3051-062-Q1-K1-H3

Method BLASTX
NCBI GI g1408471
BLAST score 620
E value 2.0e-64
Match length 139
% identity 83

NCBI Description (U48938) actin depolymerizing factor 1 [Arabidopsis

thaliana] >gi_3851707 (AF102173) actin depolymerizing

factor 1 [Arabidopsis thaliana]

Seq. No. 920

Contig ID 475_4.R1040

5'-most EST jC-gmst02400038f11a1

Method BLASTX
NCBI GI g1408471
BLAST score 630
E value 1.0e-65
Match length 139
% identity 84

NCBI Description (U48938) actin depolymerizing factor 1 [Arabidopsis thaliana] >gi_3851707 (AF102173) actin depolymerizing

factor 1 [Arabidopsis thaliana]

Seq. No. 921

Contig ID 475_5.R1040

5'-most EST LIB3087-010-Q1-K1-H12

Method BLASTX
NCBI GI g4185511
BLAST score 410
E value 4.0e-40
Match length 90
% identity 86

NCBI Description (AF102822) actin depolymerizing factor 4 [Arabidopsis

thaliana]

Seq. No. 922

Contig ID 475_6.R1040

5'-most EST uC-gmropic057g10b1

Method BLASTN
NCBI GI g1399379
BLAST score 325
E value 0.0e+00



Match length 459 % identity 95

NCBI Description Glycine max S-adenosyl-L-methionine:delta24-sterol-C-

methyltransferase mRNA, complete cds

Seq. No. 923

Contig ID 475_9.R1040 5'-most EST crh700854308.h1

Method BLASTN
NCBI GI g1399379
BLAST score 184
E value 5.0e-99
Match length 200
% identity 98

NCBI Description Glycine max S-adenosyl-L-methionine:delta24-sterol-C-

methyltransferase mRNA, complete cds

Seq. No. 924

Contig ID 476 1.R1040

5'-most EST LIB3051-085-Q1-K1-H9

Method BLASTN
NCBI GI g1236950
BLAST score 620
E value 0.0e+00
Match length 717
% identity 97

NCBI Description Glycine max nucleoside diphosphate kinase mRNA, complete

cds

Seq. No. 925

Contig ID 476 2.R1040

5'-most EST LIB3049-042-Q1-E1-A12

Method BLASTN
NCBI GI g1236950
BLAST score 372
E value 0.0e+00
Match length 609
% identity 92

NCBI Description Glycine max nucleoside diphosphate kinase mRNA, complete

cds

Seq. No. 926

Contig ID 476_3.R1040 5'-most EST ncj700981096.h1

Method BLASTN
NCBI GI g1236950
BLAST score 184
E value 3.0e-99
Match length 271
% identity 96

NCBI Description Glycine max nucleoside diphosphate kinase mRNA, complete

cds

Seq. No. 927

Contig ID 476_4.R1040 5'-most EST jsh701070081.h2

Method BLASTN



NCBI GI g1236950 BLAST score 326 E value 0.0e+00 Match length 342 % identity 99

NCBI Description Glycine max nucleoside diphosphate kinase mRNA, complete

cds

Seq. No. 928

Contig ID 477_1.R1040

5'-most EST LIB3109-035-Q1-K4-B1

Method BLASTN
NCBI GI g1277163
BLAST score 449
E value 0.0e+00
Match length 467
% identity 99

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 929

Contig ID 478_1.R1040 5'-most EST leu701154606.h1

Method BLASTN
NCBI GI g1277165
BLAST score 221
E value 1.0e-121
Match length 488
% identity 90

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 930

Contig ID 478_2.R1040 5'-most EST leu701147147.h1

Method BLASTN
NCBI GI g1277165
BLAST score 198
E value 1.0e-107
Match length 332
% identity 95

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 931

Contig ID 478 3.R1040

5'-most EST LIB3106-070-P1-K1-A4

Method BLASTN
NCBI GI g1277165
BLAST score 315
E value 1.0e-177
Match length 554
% identity 92

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No.

932

Contig ID 478_4.R1040

5'-most EST LIB3028-036-Q1-B1-A12

Method BLASTN NCBI GI g1277165



BLAST score 206 E value 1.0e-112 Match length 365 % identity 96

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 933

Contig ID 478_6.R1040

5'-most EST LIB3073-003-Q1-K1-E2

Method BLASTN
NCBI GI g1277165
BLAST score 103
E value 2.0e-50
Match length 227
% identity 92

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 934

Contig ID 479 1.R1040 5'-most EST ncj700988504.h1

Method BLASTN
NCBI GI g1277167
BLAST score 425
E value 0.0e+00
Match length 455
% identity 98

NCBI Description Glycine max cysteine proteinase inhibitor mRNA, partial cds

Seq. No. 935

Contig ID 480 2.R1040

5'-most EST LIB3051-019-Q1-E1-H7

Method BLASTN
NCBI GI g310560
BLAST score 63
E value 8.0e-27
Match length 127
% identity 98

NCBI Description Soybean ascorbate peroxidase mRNA, complete cds

Seq. No. 936

Contig ID 480 3.R1040

5'-most EST jC-gmst02400069e04a1

Method BLASTX
NCBI GI g1679794
BLAST score 415
E value 3.0e-53
Match length 372
% identity 37

NCBI Description (U77627) Allele: hi2 [Danio rerio]

Seq. No. 937

Contig ID 480_4.R1040

5'-most EST LIB3170-046-Q1-J1-G3

Method BLASTN
NCBI GI g1336081
BLAST score 78
E value 2.0e-35



Match length 102 % identity 94

NCBI Description Glycine max var. Century ascorbate peroxidase 2 (APx2)

mRNA, complete cds

Seq. No. 938

Contig ID 480 5.R1040

5'-most EST LIB3093-046-Q1-K1-H1

Method BLASTN
NCBI GI g2055227
BLAST score 44
E value 3.0e-15
Match length 111
% identity 95

NCBI Description Glycine max mRNA for SRC1, complete cds

Seq. No. 939

Contig ID 480 6.R1040

5'-most EST LIB3052-013-Q1-N1-H3

Method BLASTN
NCBI GI g310560
BLAST score 893
E value 0.0e+00
Match length 1043
% identity 98

NCBI Description Soybean ascorbate peroxidase mRNA, complete cds

Seq. No. 940

Contig ID 480 7.R1040

5'-most EST LIB3028-001-Q1-B1-F10

Method BLASTN
NCBI GI g1336081
BLAST score 640
E value 0.0e+00
Match length 1075
% identity 93

NCBI Description Glycine max var. Century ascorbate peroxidase 2 (APx2)

mRNA, complete cds

Seq. No. 941

482_1.R1040 Contig ID g1679655 FL 5'-most EST Method BLASTN q1679655 NCBI GI BLAST score 1156 0.0e + 00E value Match length 1180 99 % identity

NCBI Description Glycine max metalloproteinase mRNA, complete cds

Seq. No. 942

Contig ID 483_1.R1040 5'-most EST zsg701121157.h1

Method BLASTX
NCBI GI g1679658
BLAST score 1661
E value 0.0e+00



```
Match length 342
% identity 93
NCBI Description (U63726) gamma glutamyl hydrolase [Glycine max]
```

 Seq. No.
 943

 Contig ID
 483_2.R1040

 5'-most EST
 ncj700975264.h1

Method BLASTN
NCBI GI g1680711
BLAST score 132
E value 3.0e-68
Match length 274
% identity 98

NCBI Description Glycine max gamma glutamyl hydrolase mRNA, complete cds

Seq. No. 944

Contig ID 483_3.R1040 5'-most EST hyd700725162.h1

Method BLASTN
NCBI GI g1680711
BLAST score 136
E value 2.0e-70
Match length 188
% identity 98

NCBI Description Glycine max gamma glutamyl hydrolase mRNA, complete cds

Seq. No. 945

484 1.R1040 Contig ID 5'-most EST g1518156 FL Method BLASTX q3219787 NCBI GI 1743 BLAST score 0.0e + 00E value 345 Match length 100 % identity

NCBI Description MEIOTIC RECOMBINATION PROTEIN DMC1 HOMOLOG >gi_1518157 (U66836) RecA/Rad51/DMC1-like protein [Glycine max]

Seq. No. 946

Contig ID 486_1.R1040

5'-most EST LIB3049-048-Q1-E1-H6

Method BLASTN
NCBI GI g1575726
BLAST score 811
E value 0.0e+00
Match length 967
% identity 96

NCBI Description Glycine max 14-3-3 related protein SGF14B mRNA, partial cds

Seq. No. 947

Contig ID 486_2.R1040

5'-most EST LIB3106-069-P1-K1-H4

Method BLASTN
NCBI GI g169069
BLAST score 256
E value 1.0e-141
Match length 472



% identity 89

NCBI Description Pea Cu-Zn superoxide dismutase mRNA, complete cds

Seq. No. 948

Contig ID 486_5.R1040 5'-most EST pxt700946009.h1

Method BLASTN
NCBI GI g3860328
BLAST score 37
E value 2.0e-11
Match length 139
% identity 84

NCBI Description Cicer arietinum mRNA for superoxide dismutase

Seq. No. 949

Contig ID 486_6.R1040

5'-most EST jC-gmle01810004h01a1

Method BLASTN
NCBI GI g169069
BLAST score 93
E value 1.0e-44
Match length 289
% identity 83

NCBI Description Pea Cu-Zn superoxide dismutase mRNA, complete cds

Seq. No. 950

Contig ID 486 8.R1040

5'-most EST fC-gmf1700905524r4

Method BLASTN
NCBI GI g169069
BLAST score 57
E value 2.0e-23
Match length 85
% identity 93

NCBI Description Pea Cu-Zn superoxide dismutase mRNA, complete cds

Seq. No. 951

Contig ID 488_1.R1040 5'-most EST bth700846059.h1

Method BLASTN
NCBI GI g1575730
BLAST score 1079
E value 0.0e+00
Match length 1130
% identity 99

NCBI Description Glycine max 14-3-3 related protein SGF14D mRNA, complete

cds

Seq. No. 952

Contig ID 488_2.R1040

5'-most EST jC-gmf102220057d03a1

Method BLASTN
NCBI GI g1575730
BLAST score 233
E value 1.0e-128
Match length 257
% identity 98

248



NCBI Description Glycine max 14-3-3 related protein SGF14D mRNA, complete cds

Seq. No. 953

Contig ID 489 2.R1040 5'-most EST q5510318 BLASTN Method q1619904 NCBI GI BLAST score 151 2.0e-79 E value Match length 290 % identity 89

NCBI Description Glycine max thiol protease isoform A mRNA, partial cds

Seq. No. 954

Contig ID 489 3.R1040

5'-most EST LIB3106-070-P1-K1-D9

Method BLASTX
NCBI GI g1620986
BLAST score 327
E value 3.0e-30
Match length 78
% identity 79

NCBI Description (Y08858) 40S ribosomal protein S17 [Nicotiana

plumbaginifolia]

Seq. No. 955

Contig ID 489 4.R1040

5'-most EST jC-gmro02910023g10d1

Method BLASTN
NCBI GI g2511690
BLAST score 131
E value 2.0e-67
Match length 325
% identity 89

NCBI Description Phaseolus vulgaris Moldavian encoding cysteine proteinase

precursor (clone cp41)

Seq. No. 956

Contig ID 489_5.R1040 5'-most EST leu701156563.h1

Seq. No. 957

Contig ID 489_6.R1040 5'-most EST zzp700830691.h1

Method BLASTX
NCBI GI g1172874
BLAST score 236
E value 3.0e-41
Match length 215
% identity 41

NCBI Description DEHYDRATION-RESPONSIVE PROTEIN RD22 PRECURSOR

>gi_479589_pir__S34823 dehydration-induced protein RD22 Arabidopsis thaliana >gi_391608_dbj_BAA01546_ (D10703) rd22
[Arabidopsis thaliana] >gi_447134_prf__1913421A rd22 gene

[Arabidopsis thaliana]



```
958
Seq. No.
                  489 7.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy276g11b1
Method
                  BLASTX
NCBI GI
                  g2511691
                  539
BLAST score
                  1.0e-129
E value
                  330
Match length
                  74
% identity
                  (Z99953) cysteine proteinase precursor [Phaseolus vulgaris]
NCBI Description
                  959
Seq. No.
                   489 8.R1040
Contig ID
5'-most EST
                  uC-gmropic034b07b1
                  BLASTN
Method
                  g2257597
NCBI GI
                   436
BLAST score
                  0.0e + 00
E value
                   684
Match length
                   91
% identity
                  Robinia pseudoacacia mRNA for phosphoglycerate kinase,
NCBI Description
                  partial cds
                   960
Seq. No.
Contig ID
                   489 9.R1040
                   LIB3138-009-Q1-N1-B11
5'-most EST
Method
                   BLASTN
                   g1619904
NCBI GI
BLAST score
                   827
                   0.0e + 00
E value
                   1190
Match length
% identity
NCBI Description Glycine max thiol protease isoform A mRNA, partial cds
                   961
Seq. No.
                   491 1.R1040
Contig ID
5'-most EST
                   LIB3093-002-Q1-K1-E6
                   BLASTN
Method
                   q456713
NCBI GI
BLAST score
                   1118
E value
                   0.0e + 00
Match length
                   1160
                   100
% identity
NCBI Description Glycine max gene for ubiquitin, complete cds
                   962
Seq. No.
                   491 2.R1040
Contig ID
5'-most EST
                   LIB3093-036-Q1-K1-E8
```

Method BLASTN NCBI GI q303900 BLAST score 680 0.0e+00E value 755 Match length 46 % identity

NCBI Description Soybean gene for ubiquitin, complete cds

Seq. No. 963



```
491 4.R1040
Contig ID
                  rca701000445.hl
5'-most EST
                  BLASTN
Method
                  g456713
NCBI GI
                  160
BLAST score
                  7.0e-85
E value
                  196
Match length
                  72
% identity
NCBI Description Glycine max gene for ubiquitin, complete cds
                  964
Seq. No.
                  491 5.R1040
Contig ID
                  LIB3106-092-Q1-K1-E11
5'-most EST
                  BLASTN
Method
NCBI GI
                  q456713
                  209
BLAST score
                   1.0e-114
E value
                   257
Match length
% identity
NCBI Description Glycine max gene for ubiquitin, complete cds
                   965
Seq. No.
                   492 1.R1040
Contig ID
                   leu701145206.hl
5'-most EST
                   BLASTN
Method
                   g18781
NCBI GI
                   403
BLAST score
                   0.0e + 00
E value
                   403
Match length
                   100
% identity
NCBI Description G.max win gene encoding wound-induced protein
                   966
Seq. No.
                   492_2.R1040
Contig ID
                   fC-gmro700848126a1
5'-most EST
                   BLASTN
Method
                   q18781
NCBI GI
                   177
BLAST score
                   1.0e-94
E value
Match length
                   249
% identity
NCBI Description G.max win gene encoding wound-induced protein
Seq. No.
                   493 1.R1040
Contig ID
                   g872115 FL
5'-most EST
                   BLASTX
Method
NCBI GI
                   g872116
BLAST score
                   2461
                   0.0e+00
E value
Match length
                   569
                   85
 % identity
                   (X79770) sti (stress inducible protein) [Glycine max]
NCBI Description
                   968
```

Seq. No.

494 1.R1040 Contig ID pxt700946118.hl 5'-most EST



Method BLASTX
NCBI GI g1076498
BLAST score 1980
E value 0.0e+00
Match length 389
% identity 96

NCBI Description zinc-finger protein (C-terminal) - soybean

>gi 558543 emb CAA85320 (Z36749) C-terminal zinc-finger

[Glycine max]

Seq. No. 969

Contig ID 497 1.R1040

5'-most EST LIB3139-045-P1-N1-G1

Method BLASTX
NCBI GI g129585
BLAST score 3399
E value 0.0e+00
Match length 706
% identity 93

NCBI Description PHENYLALANINE AMMONIA-LYASE CLASS II >gi_81877_pir__S04127

phenylalanine ammonia-lyase (EC 4.3.1.5) class II - kidney bean >gi 228614 prf 1807329A Phe ammonia lyase [Phaseolus

vulgaris]

Seq. No. 970

Contig ID 497_2.R1040

5'-most EST fC-gmro700807584e1

Method BLASTN
NCBI GI g18376
BLAST score 572
E value 0.0e+00
Match length 592
% identity 99

NCBI Description Glycine max PAL1 gene for phenylalanine ammonia lyase (EC

4.3.1.5)

Seq. No. 971

Contig ID 497_3.R1040

5'-most EST uC-gmrominsoy056b10b1

Method BLASTX
NCBI GI g266731
BLAST score 622
E value 7.0e-65
Match length 144
% identity 85

NCBI Description PHENYLALANINE AMMONIA-LYASE 1 >gi_282927_pir_S25303

phenylalanine ammonia-lyase (EC 4.3.1.5) - garden pea

>gi_217980_dbj_BAA00885_ (D10001) phenylalanine

ammonia-lyase [Pisum sativum] >gi_217982_dbj_BAA00886_(D10002) phenylalanine ammonia-lyase [Pisum sativum]

Seq. No. 972

Contig ID 498 1.R1040

5'-most EST uC-gmrominsoy123d08b1

Method BLASTX NCBI GI g231799 BLAST score 1957



```
0.0e + 00
E value
                  379
Match length
                  99
% identity
                  CHALCONE SYNTHASE 7 (NARINGENIN-CHALCONE SYNTHASE 7)
NCBI Description
                  >gi_478407_pir__JQ2250 naringenin-chalcone synthase (EC
                  2.3.1.74) - soybean >gi_169937 (M98871) chalcone synthase
                   [Glycine max]
                   973
Seq. No.
                   498 2.R1040
Contig ID
                   g4300400
5'-most EST
                   BLASTN
Method
                   g169936
NCBI GI
BLAST score
                   860
                   0.0e + 00
E value
                   999
Match length
                   97
% identity
                  Glycine max chalcone synthase (chs7) gene, complete cds
NCBI Description
                   974
Seq. No.
                   498_3.R1040
Contig ID
                   uC-gmropic108h09b1
5'-most EST
                   BLASTN
Method
                   g218012
NCBI GI
BLAST score
                   385
                   0.0e + 00
E value
                   513
Match length
                   95
% identity
NCBI Description Pueraria lobata mRNA for chalocone synthase, complete cds
                   975
Seq. No.
                   498 6.R1040
Contig ID
                   uC-qmronoir028c02b1
5'-most EST
Method
                   BLASTN
                   q169936
NCBI GI
BLAST score
                   301
E value
                   1.0e-169
                   334
Match length
% identity
                   99
NCBI Description Glycine max chalcone synthase (chs7) gene, complete cds
                   976
Seq. No.
                   499 1.R1040
Contig ID
5'-most EST
                   g545762 FL
Method
                   BLASTN
                   g545762
NCBI GI
BLAST score
                   518
                   0.0e + 00
E value
                   526
Match length
                   100
% identity
```

Eu4=urease {clone 5E5} [Glycine max=soybeans, cv. Dare, NCBI Description

mid-maturation embryos, mRNA Partial, 540 nt]

977 Seq. No.

500 1.R1040 Contig ID

LIB3170-038-Q1-K1-E6 5'-most EST

BLASTX Method



q3309269 NCBI GI 1810 BLAST score 0.0e + 00E value 501 Match length % identity 97

(AF074940) ferric leghemoglobin reductase-2 precursor NCBI Description

[Glycine max]

978 Seq. No.

500 3.R1040 Contig ID

LIB3087-008-Q1-K1-B10 5'-most EST

BLASTN Method g546359 NCBI GI 279 BLAST score 1.0e-155 E value Match length 327 % identity 96

ferric leghemoglobin reductase [Glycine max=soybeans, Merr, NCBI Description

nodules, mRNA, 1740 nt]

979 Seq. No.

504 1.R1040 Contig ID

uC-gmflminsoy062a07b1 5'-most EST

BLASTX Method g1706522 NCBI GI BLAST score 2613 0.0e + 00E value 521 Match length 94 % identity

DIHYDROFOLATE REDUCTASE / THYMIDYLATE SYNTHASE (DHFR-TS) NCBI Description

>gi_1362048_pir__S55683 dihydrofolate reductase-thymidylate
synthase - soybean >gi_999190_bbs_166832 (S78087)

dihydrofolate reductase-thymidylate synthase,

DHFR-TS=bifunctional enzyme {EC 1.5.1.3} [Glycine max,

seedling, Peptide, 530 aa] [Glycine max] >gi 1096144_prf__2111237A dihydrofolate reductase-thymidylate synthase [Glycine max]

980 Seq. No.

504 2.R1040 Contig ID

 $fC-\overline{g}mst700656887r1$ 5'-most EST

BLASTN Method q999189 NCBI GI 146 BLAST score 3.0e-76 E value 166 Match length 97 % identity

dihydrofolate reductase-thymidylate synthase=bifunctional NCBI Description

enzyme [Glycine max, seedling, mRNA, 1794 nt]

981 Seq. No.

504 4.R1040 Contig ID

fC-qmro7000746445a1 5'-most EST

BLASTN Method g999189 NCBI GI 169 BLAST score 4.0e-90 E value



257 Match length 91 % identity

dihydrofolate reductase-thymidylate synthase=bifunctional NCBI Description

enzyme [Glycine max, seedling, mRNA, 1794 nt]

982 Seq. No.

507 1.R1040 Contig ID

LIB3065-026-Q1-N1-B7 5'-most EST

BLASTN Method NCBI GI g498167 525 BLAST score 0.0e+00E value 545 Match length 99 % identity

NCBI Description Soybean mRNA for leginsulin, complete cds

983 Seq. No.

517 1.R1040 Contig ID g218259 FL 5'-most EST BLASTX Method NCBI GI g730164 2436 BLAST score 0.0e+00E value 482 Match length

100 % identity EARLY NODULIN 70 >gi_486678_pir__S34800 sulfate transport NCBI Description

protein homolog (clone GmN70) - soybean

>gi_218260_dbj_BAA02723_ (D13505) early nodulin [Glycine

max] >gi_4 $\overline{47137}$ _prf__19 $\overline{13422C}$ nodulin [Glycine max]

984 Seq. No.

518 2.R1040 Contig ID

LIB3040-044-Q1-E1-C8 5'-most EST

BLASTN Method g218261 NCBI GI 121 BLAST score 2.0e-61 E value 229 Match length % identity

Soybean mRNA for early nodulin NCBI Description

985 Seq. No.

521 1.R1040 Contig ID seb700650013.hl 5'-most EST

BLASTX Method q485495 NCBI GI 2326 BLAST score 0.0e + 00E value 463 Match length 97 % identity NCBI Description

aspartate transaminase (EC 2.6.1.1) AAT5 precursor soybean >gi_169915 (L09702) aspartate aminotransferase [Glycine max] >gi_300419_bbs_132011 (S60967) aspartate aminotransferase isozyme 5, AAT5 {EC 2.6.1.1} [Glycine max=soybeans, cv. Century, Peptide Chloroplast, 463 aa]

[Glycine max]



```
986
Seq. No.
                  521 2.R1040
Contig ID
                  LIB3049-021-Q1-E1-C8
5'-most EST
                  BLASTN
Method
                  q300418
NCBI GI
                  290
BLAST score
                  1.0e-162
E value
                  342
Match length
% identity
                 aspartate aminotransferase isozyme 5 [Glycine max=soybeans,
NCBI Description
                  cv. Century, mRNA, 1755 nt]
                  987
Seq. No.
                  521 3.R1040
Contig ID
                  gsv701055975.hl
5'-most EST
                  BLASTN
Method
                  q169914
NCBI GI
                  123
BLAST score
                   7.0e-63
E value
                   123
Match length
                   100
% identity
NCBI Description Glycine max (clone pSAT17) aspartate aminotransferase mRNA,
                   complete cds
                   988
Seq. No.
                   522_1.R1040
Contig ID
                   g169918 FL
5'-most EST
                   BLASTN
Method
                   g169918
NCBI GI
                   1735
BLAST score
                   0.0e + 00
E value
                   1924
Match length
                   99
% identity
NCBI Description Soybean auxin-regulated protein (Aux22) mRNA, complete cds
                   989
Seq. No.
                   522 2.R1040
Contig ID
                   rlr700901873.hl
5'-most EST
                   BLASTN
Method
                   g169918
NCBI GI
                   144
BLAST score
                   3.0e-75
E value
                   316
Match length
                   88
% identity
NCBI Description Soybean auxin-regulated protein (Aux22) mRNA, complete cds
Seq. No.
                   990
                   523 1.R1040
Contig ID
                   g169920 FL
 5'-most EST
                   BLASTN
Method
NCBI GI
                   g169920
BLAST score
                   1780
                    0.0e + 00
E value
```

NCBI Description Soybean auxin-regulated protein (Aux28) mRNA, complete cds

1784

100

Match length

% identity



991 Seq. No. 523 2.R1040 Contig ID leu701151952.h1 5'-most EST Method BLASTN g169920 NCBI GI 213 BLAST score 1.0e-116 E value 531 Match length 87 % identity Soybean auxin-regulated protein (Aux28) mRNA, complete cds NCBI Description 992 Seq. No. 523 3.R1040 Contig ID jC-gmf102220103g11d1 5'-most EST Method BLASTN NCBI GI g169920 103 BLAST score 1.0e-50 E value 285 Match length 87 % identity NCBI Description Soybean auxin-regulated protein (Aux28) mRNA, complete cds 993 Seq. No. 523 4.R1040 Contig ID seb700653443.h1 5'-most EST BLASTX Method g114734 NCBI GI BLAST score 192 2.0e-14 E value 113 Match length 47 % identity AUXIN-INDUCED PROTEIN AUX28 >gi_81759 pir_ A28993 aux28 NCBI Description protein - soybean >gi_169921 (JO3919) auxin-regulated protein (Aux28) [Glycine max] 994 Seq. No. 524_1.R1040 Contig ID zsg701127613.h1 5'-most EST BLASTX Method q1706546 NCBI GI 1791

Method BLASTX
NCBI GI g170654
BLAST score 1791
E value 0.0e+00
Match length 347
% identity 100

NCBI Description GLUCAN ENDO-1,3-BETA-GLUCOSIDASE PRECURSOR

((1->3)-BETA-GLUCAN ENDOHYDROLASE) ((1->3)-BETA-GLUCANASE)

(BETA-1,3-ENDOGLUCANASE) >gi_169923 (M37753)

beta-1,3-endoglucanase (EC 3.2.1.39) [Glycine max]

Seq. No. 995

Contig ID 524 2.R1040 5'-most EST zzp700832021.h1

Method BLASTX
NCBI GI g1403675
BLAST score 760
E value 7.0e-81
Match length 192



% identity NCBI Description (U41323) beta-1,3-glucanase [Glycine max] 996 Seq. No. 524 3.R1040 Contig ID jC-gmst02400051f12a1 5'-most EST BLASTN Method g1403674 NCBI GI BLAST score 273 1.0e-152 E value 449 Match length 90 % identity NCBI Description Glycine max beta-1,3-glucanase (SGN1) gene, complete cds 997 Seq. No. 524 4.R1040 Contig ID LIB3139-048-P1-N1-H7 5'-most EST Method BLASTX NCBI GI q1403675 528 BLAST score 7.0e-54 E value 139 Match length 75 % identity (U41323) beta-1,3-glucanase [Glycine max] NCBI Description 998 Seq. No. 528 1.R1040 Contig ID uC-gmrominsoy264f06b1 5'-most EST Method BLASTX NCBI GI g116054 BLAST score 2564 E value 0.0e+00Match length 492 % identity 100 NCBI Description CALCIUM-DEPENDENT PROTEIN KINASE SK5 (CDPK) >gi_280393_pir__A43713 calcium-dependent protein kinase (EC 2.7.1.-) - soybean >gi 169931 (M64987) Glycine max calcium dependent protein kinase mRNA. [Glycine max] 999 Seq. No. 528 2.R1040 Contig ID fua701040310.h1 5'-most EST BLASTN Method g169930 NCBI GI BLAST score 279 1.0e-155 E value 423 Match length 92 % identity NCBI Description Glycine max calcium dependent protein kinase mRNA

Seq. No.

1000

Contig ID 528 3.R1040

5'-most EST fC-gmse700674513d3

Method BLASTN
NCBI GI g169930
BLAST score 187
E value 1.0e-101

258

一切地震力.



Match length % identity 89

NCBI Description Glycine max calcium dependent protein kinase mRNA

Seq. No.

1001

Contig ID

531 1.R1040

5'-most EST Method

jC-gmst02400047f05a1

NCBI GI

BLASTX g3096935

BLAST score E value Match length

8.0e-15

% identity

102 42

195

NCBI Description

(AL023094) putative protein [Arabidopsis thaliana]

Seq. No.

1002

Contig ID 5'-most EST 531 2.R1040 g5752719

Method NCBI GI BLAST score BLASTX q1617206 275

E value Match length % identity

4.0e-24 129 48

NCBI Description

(Z72489) CP12 [Pisum sativum]

Seq. No.

1003

Contig ID

531 3.R1040

5'-most EST

LIB3170-045-Q1-K1-E4 BLASTX

Method NCBI GI BLAST score

g3334150 1870

E value Match length 0.0e + 00421

% identity

90

NCBI Description

MAGNESIUM-CHELATASE SUBUNIT CHLI PRECURSOR

(MG-PROTOPORPHYRIN IX CHELATASE) >gi_2129847_pir__JC4312 chlorophyll magnesium chelatase (EC $\overline{4.99.-.-}$) - \overline{soy} bean chloroplast >gi_1732469_dbj_BAA08291_ (D45857) Mg chelatase subunit (46 kD) [Glycine max]

Seq. No.

1004

Contig ID

531 4.R1040

5'-most EST

LIB3028-039-Q1-B2-B2

Method NCBI GI BLASTN

BLAST score

q1732468

E value

82

Match length

5.0e-38 86

% identity NCBI Description

99 Soybean mRNA for Mg chelatase subunit (46 kD), complete cds

Seq. No.

1005

Contig ID 5'-most EST 531 5.R1040

g5510205

Seq. No.

1006



```
531 6.R1040
Contig ID
                   LIB3107-014-Q1-K1-D9
5'-most EST
Method
                   BLASTN
NCBI GI
                   g1732468
BLAST score
                   133
                   3.0e-68
E value
                   439
Match length
                   91
% identity
                  Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
NCBI Description
                   1007
Seq. No.
                   531 7.R1040
Contig ID
                   LIB3074-011-Q1-E1-E12
5'-most EST
Method
                   BLASTN
NCBI GI
                   g1732468
BLAST score
                   82
                   3.0e-38
E value
                   86
Match length
                   99
% identity
                  Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
NCBI Description
                   1008
Seq. No.
                   531 8.R1040
Contig ID
                   g5677965
5'-most EST
                   BLASTX
Method
                   g1617206
NCBI GI
BLAST score
                   244
                   1.0e-20
E value
                   79
Match length
                   57
% identity
                  (Z72489) CP12 [Pisum sativum]
NCBI Description
                   1009
Seq. No.
                   531_9.R1040
Contig ID
5'-most EST
                   uC-gmropic110e07b1
                   BLASTN
Method
NCBI GI
                   g1732468
                   189
BLAST score
                   1.0e-102
E value
Match length
                   257
% identity
NCBI Description Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
                   1010
Seq. No.
Contig ID
                   531_10.R1040
5'-most EST
                   g5752984
Method
                   BLASTN
NCBI GI
                   g1732468
BLAST score
                   144
E value
                   4.0e-75
Match length
                   432
% identity
                   98
NCBI Description Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
```

Seq. No. 1011

Contig ID 531 14.R1040 5'-most EST ssr700557049.h1



```
BLASTN
Method
                  q1732468
NCBI GI
BLAST score
                  180
                  7.0e-97
E value
Match length
                  244
                  93
% identity
NCBI Description Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
Seq. No.
                  1012
                  534_1.R1040
Contig ID
                   jC-gmf102220070f11a1
5'-most EST
                  BLASTX
Method
                   g1351410
NCBI GI
BLAST score
                   2477
E value
                   0.0e+00
Match length
                   495
                   94
% identity
                  VACUOLAR PROCESSING ENZYME PRECURSOR (VPE)
NCBI Description
                   >gi 511938 dbj_BAA06030_ (D28876) cysteine proteinase
                   [Glycine max]
                   1013
Seq. No.
                   534 2.R1040
Contig ID
                   LIB3107-004-Q1-K1-D4
5'-most EST
                   BLASTN
Method
                   g511937
NCBI GI
                   244
BLAST score
                   1.0e-135
E value
                   326
Match length
                   94
% identity
                   Soybean mRNA for cysteine proteinase, complete cds
NCBI Description
                   1014
Seq. No.
                   534_3.R1040
Contig ID
                   uC-gmropic103c05b1
5'-most EST
Method
                   BLASTN
NCBI GI
                   q511937
                   267
BLAST score
                   1.0e-148
E value
Match length
                   279
% identity
                   99
                   Soybean mRNA for cysteine proteinase, complete cds
NCBI Description
                                                             ٠ ٢.
                   1015
Seq. No.
                   535 1.R1040
Contig ID
5'-most EST
                   seb700653014.hl
Method
                   BLASTX
                   g2494415
NCBI GI
                   1914
BLAST score
                   0.0e + 00
E value
                   394
Match length
                   94
% identity
                   FRUCTOSE-1, 6-BISPHOSPHATASE, CHLOROPLAST PRECURSOR
NCBI Description
                   (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)
```

max]

>gi 515747 (L34841) fructose-1,6-bisphosphatase [Glycine

```
Seq. No.
                   1016
Contig ID
                   535 2.R1040
                  vzy700755037.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                   q20716
BLAST score
                   155
E value
                   9.0e-82
                  294
Match length
                   88
% identity
NCBI Description P.sativum mRNA for fructose 1,6 biphosphatase
                  1017
Seq. No.
Contig ID
                   535 3.R1040
5'-most EST
                   fde700872212.h1
Method
                  BLASTN
                   g515746
NCBI GI
BLAST score
                   166
E value
                   2.0e-88
                   187
Match length
% identity
                   97
                  Soybean chloroplast fructose-1,6-bisphosphatase (FBP) mRNA,
NCBI Description
                   complete cds
Seq. No.
                   1018
                   537 1.R1040
Contig ID
                  LIB3040-043-Q1-E1-C11
5'-most EST
                   BLASTN
Method
NCBI GI
                   a548320
BLAST score
                   944
E value
                   0.0e + 00
Match length
                   984
% identity
                   99
NCBI Description
                   complete cds
```

Glycine max dihydrodipicolinate synthase (dapA) mRNA,

Seq. No. 1019

537 2.R1040 Contig ID 5'-most EST ncj700981542.h1

Method BLASTN g548320 NCBI GI BLAST score 126 2.0e-64 E value 138 Match length 98 % identity

NCBI Description Glycine max dihydrodipicolinate synthase (dapA) mRNA,

complete cds

Seq. No. 1020

539 1.R1040 Contig ID

5'-most EST jC-gmf102220138d09a1

BLASTN Method NCBI GI q457569 958 BLAST score 0.0e+00E value Match length 962 % identity 100

NCBI Description Soybean mRNA for endo-xyloglucan transferase, partial cds



```
1021
Seq. No.
                   539 2.R1040
Contig ID
                   leu701153112.h1
5'-most EST
Method
                   BLASTN
                   g457569
NCBI GI
BLAST score
                   373
                   0.0e+00
E value
                   780
Match length
% identity
                   88
NCBI Description
                  Soybean mRNA for endo-xyloglucan transferase, partial cds
Seq. No.
                   1022
                   539 3.R1040
Contig ID
5'-most EST
                   LIB3028-050-Q1-B1-A12
Method
                   BLASTX
NCBI GI
                   g2813966
BLAST score
                   234
E value
                   6.0e-19
Match length
                   200
% identity
NCBI Description
                   (Z97053) placental protein Diff33 [Homo sapiens]
                   >gi 4545219 gb AAD22448.1 (AF112227) TDE homolog [Homo
                   sapiens]
Seq. No.
                   1023
                   539 4.R1040
Contig ID
5'-most EST
                   ncj700981722.h1
                   BLASTX
Method
NCBI GI
                   g2813966
BLAST score
                   255
E value
                   2.0e-21
Match length
                   150
                   35
% identity
                   (Z97053) placental protein Diff33 [Homo sapiens]
NCBI Description
                   >gi 4545219 gb AAD22448.1 (AF112227) TDE homolog [Homo
                   sapiens]
                   1024
Seq. No.
                   539 5.R1040
Contig ID
5'-most EST
                   leu701144894.h1
Method
                   BLASTN
NCBI GI
                   g457572
BLAST score
                   631
E value
                   0.0e + 00
Match length
                   905
% identity
                   93
NCBI Description
                  V.angularis mRNA for endo-xyloglucan transferase, complete
                   cds
Seq. No.
                   1025
```

Contig ID 539_6.R1040 5'-most EST zhf700959501.h1

Seq. No. 1026

Contig ID 539_8.R1040

5'-most EST jC-gmst02400018d10d1

Match length

% identity

1044

100

NCBI Description Glycine max ferritin mRNA, complete cds



```
Method
NCBI GI
                   q457569
BLAST score
                   250
E value
                   1.0e-138
Match length
                   250
% identity
                   100
                  Soybean mRNA for endo-xyloglucan transferase, partial cds
NCBI Description
Seq. No.
                   1027
Contig ID
                   539 9.R1040
5'-most EST
                  LIB3028-006-Q1-B1-C3
Method
                  BLASTN
NCBI GI
                   q457569
BLAST score
                   206
E value
                   1.0e-112
Match length
                   298
% identity
                  Soybean mRNA for endo-xyloglucan transferase, partial cds
NCBI Description
Seq. No.
                   1028
Contig ID
                   539 11.R1040
5'-most EST
                   zzp700832821.h1
Seq. No.
                   1029
Contig ID
                   539 12.R1040
5'-most EST
                  LIB3028-024-Q1-B1-D8
Method
                  BLASTN
NCBI GI
                  q457569
BLAST score
                  273
                  1.0e-152
E value
Match length
                  297
% identity
                  98
NCBI Description Soybean mRNA for endo-xyloglucan transferase, partial cds
Seq. No.
                  1030
Contig ID
                  539 13.R1040
5'-most EST
                  jC-gmf102220082h12a1
Method
                  BLASTN
NCBI GI
                  g457572
BLAST score
                  387
E value
                  0.0e+00
Match length
                  666
% identity
                  V.angularis mRNA for endo-xyloglucan transferase, complete
NCBI Description
                  cds
Seq. No.
                  1031
Contig ID
                  540 1.R1040
                  pxt700944790.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g169952
BLAST score
                  1024
E value
                  0.0e+00
```



Seq. No. 1032

Contig ID 540_2.R1040

5'-most EST LIB3170-037-Q1-K1-H5

Method BLASTN
NCBI GI g170077
BLAST score 287
E value 1.0e-160
Match length 311
% identity 98

NCBI Description Soybean ferritin (SOF-H2) mRNA, partial cds

Seq. No. 1033

Contig ID 542 1.R1040

5'-most EST LIB3138-079-P1-N1-B2

Method BLASTN
NCBI GI g169954
BLAST score 1048
E value 0.0e+00
Match length 1048
% identity 100

NCBI Description Glycine max iron superoxide dismutase (FeSOD) mRNA,

complete cds

Seq. No. 1034

Contig ID 542_3.R1040 5'-most EST wrg700792105.h1

Method BLASTX
NCBI GI g134646
BLAST score 220
E value 7.0e-18
Match length 49
% identity 73

NCBI Description SUPEROXIDE DISMUTASE, CHLOROPLAST PRECURSOR (FE) >gi 169955

(M64267) Fe-superoxide dismutase [Glycine max]

>gi_228415_prf__1803527A Fe superoxide dismutase [Glycine

max]

Seq. No. 1035

Contig ID 542_5.R1040 5'-most EST rca701001755.h1

Method BLASTN
NCBI GI g169954
BLAST score 147
E value 4.0e-77
Match length 213
% identity 92

NCBI Description Glycine max iron superoxide dismutase (FeSOD) mRNA,

complete cds

Seq. No. 1036

Contig ID 544 1.R1040

5'-most EST uC-gmropic022d09b1

Method BLASTN
NCBI GI g169956
BLAST score 1266
E value 0.0e+00
Match length 1287



% identity NCBI Description Glycine max G-box binding factor (GBF1) mRNA, complete cds Seq. No. 1037 Contig ID 545 1.R1040 5'-most EST LIB3051-047-Q1-K1-D12 Method BLASTX NCBI GI g169959 BLAST score 1989 E value 0.0e + 00Match length 423 % identity 94

NCBI Description (L01448) G-box binding factor [Glycine max]

Seq. No. 1038 545 2.R1040 Contig ID 5'-most EST LIB3138-082-P1-N1-H1 Method BLASTX

NCBI GI g169961 BLAST score 1837 E value 0.0e + 00Match length 365 % identity 100

(L01449) G-box binding factor [Glycine max] NCBI Description

Seq. No. 1039

Contig ID 545 3.R1040

5'-most EST LIB3093-025-Q1-K1-F3

Method BLASTN NCBI GI g169958 BLAST score 191 E value 1.0e-103 Match length 358 % identity 90

NCBI Description Glycine max G-box binding factor (GBF2A) mRNA, 3' end

Seq. No. 1040 545 4.R1040 Contig ID 5'-most EST g5688140 Method BLASTX NCBI GI g1155054 BLAST score 278 E value 9.0e-44Match length 133 % identity

NCBI Description (U41817) regulator of MAT2 [Phaseolus vulgaris]

Seq. No. 1041

Contig ID 545 6.R1040

5'-most EST uC-gmropic110b11b1 Method BLASTN

NCBI GI g169958 BLAST score 47 E value 3.0e-17 Match length 118 % identity 90

NCBI Description Glycine max G-box binding factor (GBF2A) mRNA, 3' end



Seq. No. 1042

Contig ID 547 1.R1040

5'-most EST uC-gmropic074d11b1

Method BLASTX
NCBI GI g1346192
BLAST score 2517
E value 0.0e+00
Match length 525
% identity 94

NCBI Description GLUTATHIONE REDUCTASE, CHLOROPLAST PRECURSOR (GR) (GRASE) >gi_169965 (L11632) glutathione reductase [Glycine max]

Seq. No. 1043

Contig ID 547_2.R1040 5'-most EST fde700872675.h1

Method BLASTN
NCBI GI g169964
BLAST score 100
E value 7.0e-49
Match length 415
% identity 88

NCBI Description Soybean glutathione reductase (GR) mRNA, complete cds

Seq. No. 1044

Contig ID 552 1.R1040 5'-most EST q438898 FL Method BLASTX NCBI GI g1709918 BLAST score 2776 E value 0.0e+00 Match length 548 % identity 98

NCBI Description AMIDOPHOSPHORIBOSYLTRANSFERASE PRECURSOR (GLUTAMINE

PHOSPHORIBOSYLPYROPHOSPHATE AMIDOTRANSFERASE) (ATASE)

(GPAT) >gi_438899 (L23833) glutamine

phosphoribosylpyrophosphate amidotransferase [Glycine max]

Seq. No. 1045

Contig ID 552 2.R1040 5'-most EST g4303654 Method BLASTX NCBI GI g1709918 BLAST score 1084 1.0e-118 E value Match length 269 % identity 85

NCBI Description AMIDOPHOSPHORIBOSYLTRANSFERASE PRECURSOR (GLUTAMINE

PHOSPHORIBOSYLPYROPHOSPHATE AMIDOTRANSFERASE) (ATASE)

(GPAT) >gi 438899 (L23833) glutamine

phosphoribosylpyrophosphate amidotransferase [Glycine max]

Seq. No. 1046

Contig ID 553_1.R1040

5'-most EST jC-gmro02910062f11a1

Method BLASTX NCBI GI g1346105 BLAST score 1978 E value 0.0e+00 Match length 385 % identity 99

NCBI Description GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-1 SUBUNIT

(GP-ALPHA-1) >gi_439617 (L27418) G protein alpha subunit

[Glycine max]

Seq. No. 1047

Contig ID 554 1.R1040

5'-most EST LIB3049-032-Q1-E1-A12

Method BLASTX
NCBI GI g1170031
BLAST score 2323
E value 0.0e+00
Match length 466
% identity 97

NCBI Description GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE PRECURSOR (GSA)

(GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) >gi_541940_pir__JQ2263 glutamate 1-semialdehyde aminotransferase (EC 2.6.1.-) precursor - soybean >gi 310567 (L12453) glutamate 1-semialdehyde

aminotransferase [Glycine max] >gi_747968 (U20260)

glutamate 1-semialdehyde aminotransferase [Glycine max]

Seq. No. 1048

Contig ID 554_2.R1040

5'-most EST LIB3138-061-Q1-N1-G9

Method BLASTN
NCBI GI g310566
BLAST score 300
E value 1.0e-168
Match length 419
% identity 95

NCBI Description Soybean glutamate 1-semialdehyde aminotransferase mRNA,

complete cds

Seq. No. 1049

Contig ID 555 1.R1040

5'-most EST LIB3055-006-Q1-N1-B8

Method BLASTN
NCBI GI g450634
BLAST score 954
E value 0.0e+00
Match length 998
% identity 99

NCBI Description Soybean GTP binding protein mRNA, complete cds

Seq. No. 1050

Contig ID 555 2.R1040 5'-most EST g1381677 FL Method BLASTN NCBI GI q1381677 BLAST score 706 E value 0.0e+00Match length 888 % identity 95



Glycine max small GTP-binding protein (sra2) mRNA, complete NCBI Description

(in -

Seq. No.

1051

Contig ID

555 3.R1040

5'-most EST

uC-gmropic054b10b1

Method NCBI GI BLASTN q303731

BLAST score E value

456 0.0e + 00

Match length % identity

723 92

NCBI Description Pea mRNA for GTP-binding protein, complete cds

Seq. No.

1052

Contig ID 5'-most EST 555 4.R1040

Method

LIB3050-014-Q1-E1-C4 BLASTN

NCBI GI BLAST score

q1381677 336

E value Match length

0.0e + 00443

% identity

97

NCBI Description Glycine max small GTP-binding protein (sra2) mRNA, complete

cds

Seq. No.

1053

Contig ID

556 1.R1040

5'-most EST Method

LIB3109-016-Q1-K1-B2 BLASTX

NCBI GI BLAST score

q1170312 1640

E value Match length 0.0e+00379

% identity

82

NCBI Description

HOMEOBOX PROTEIN SBH1 >gi_629614_pir__S42543 hypothetical protein - soybean >gi_485406 (L13663) SBH1 [Glycine max]

Seq. No.

1054

Contig ID 5'-most EST 556 2.R1040 q5666882

Method NCBI GI

BLASTN g310568

BLAST score

279

E value Match length 1.0e-155

% identity

455

NCBI Description Glycine max homeobox-containing (Sbh1) mRNA, complete cds

Seq. No.

1055

91

Contig ID 5'-most EST

557 1.R1040 q4313302

Method NCBI GI BLAST score

BLASTN g170048 393

E value Match length 0.0e+00494



% identity 10
NCBI Description Glycine max SbPRP1 gene encoding a proline-rich protein,

complete cds

Seq. No. 1056

558 1.R1040 Contig ID 5'-most EST g4313671 BLASTN Method g347452 NCBI GI 507 BLAST score 0.0e + 00E value 876 Match length 20 % identity

NCBI Description Soybean hydroxyproline-rich glycoprotein (sbHRGP3) mRNA,

partial cds

Seq. No. 1057

Contig ID 558 2.R1040 q5606931 5'-most EST BLASTN Method g1165321 NCBI GI BLAST score 146 4.0e-76 E value Match length 560 % identity 25

NCBI Description Glycine max extensin (SbHRGP3) gene, complete cds

Seq. No. 1058

Contig ID 559 1.R1040 q4314026 5'-most EST BLASTN Method q347454 NCBI GI BLAST score 573 E value 0.0e+00Match length 736 % identity 74

NCBI Description Soybean hydroxyproline-rich glycoprotein (sbHRGP2) mRNA, 3'

end

Seq. No. 1059

Contig ID 559 2.R1040

5'-most EST LIB3107-063-Q1-K1-C2

Method BLASTN
NCBI GI g2677823
BLAST score 117
E value 1.0e-58
Match length 287
% identity 87

NCBI Description Prunus armeniaca abscisic stress ripening protein homolog

mRNA, complete cds

Seq. No. 1060

Contig ID 559_3.R1040

5'-most EST jC-gmro02910061a12a1

Method BLASTN NCBI GI g2677823

BLAST score 94



E value 4.0e-45 Match length 254 % identity 84

NCBI Description Prunus armeniaca abscisic stress ripening protein homolog

mRNA, complete cds

Seq. No. 1061

Contig ID 559_4.R1040 5'-most EST rca701000967.h1 Method BLASTN

NCBI GI g2677823 BLAST score 38 E value 9.0e-12 Match length 46 % identity 96

NCBI Description Prunus armeniaca abscisic stress ripening protein homolog

mRNA, complete cds

Seq. No. 1062

Contig ID 559_5.R1040 5'-most EST sat701009243.h1

Method BLASTN
NCBI GI g2677823
BLAST score 90
E value 6.0e-43
Match length 254
% identity 84

NCBI Description Prunus armeniaca abscisic stress ripening protein homolog

mRNA, complete cds

Seq. No. 1063

Contig ID 559 6.R1040

5'-most EST uC-gmflminsoy063c10b1

Method BLASTX
NCBI GI g4544436
BLAST score 273
E value 4.0e-24
Match length 114
% identity 50

NCBI Description (AC006592) anthocyanidin-3-glucoside rhamnosyltransferase,

3' partial [Arabidopsis thaliana]

Seq. No. 1064

Contig ID 559 9.R1040

5'-most EST jC-gmst02400033b01a1

Seq. No. 1065

Contig ID 559_11.R1040 5'-most EST g4277016

Seq. No. 1066

Contig ID 559_13.R1040 5'-most EST pmv700889430.h1

Method BLASTN
NCBI GI g2677823
BLAST score 83
E value 5.0e-39

271

134



Match length 90 % identity

NCBI Description Prunus armeniaca abscisic stress ripening protein homolog

mRNA, complete cds

1067 Seq. No.

559 19.R1040 Contig ID 5'-most EST pmv700890409.h1

BLASTN Method NCBI GI g2677823 BLAST score 48 4.0e-18 E value Match length 143

% identity 84

Prunus armeniaca abscisic stress ripening protein homolog NCBI Description

mRNA, complete cds

1068 Seq. No.

561 1.R1040 Contig ID taw700658110.h1 5'-most EST

Method BLASTX NCBI GI q1168290 BLAST score 2732 E value 0.0e+00Match length 550 % identity 95

ISOCITRATE LYASE 2 (ISOCITRASE 2) (ISOCITRATASE 2) (ICL 2) NCBI Description

>gi 349329 (L02330) glyoxysomal isocitrate lyase [Glycine

max]

1069 Seq. No.

561_2.R1040 Contig ID

5'-most EST fC-gmse700657704f5

Method BLASTX NCBI GI g1168289 BLAST score 2740 E value 0.0e + 00Match length 549 % identity 96

ISOCITRATE LYASE 1 (ISOCITRASE 1) (ISOCITRATASE 1) (ICL 1) NCBI Description

>gi 349327 (L02329) glyoxysomal isocitrate lyase [Glycine

max]

1071

1070 Seq. No.

564 1.R1040 Contig ID 5'-most EST g310579 FL Method BLASTX NCBI GI q310580 BLAST score 1761 0.0e+00E value 339 Match length 100 % identity

NCBI Description (L19360) protein kinase 2 [Glycine max]

Seq. No.

564 2.R1040 Contig ID 5'-most EST g169990 FL



```
Method
                  BLASTX
                  g1362050
NCBI GI
BLAST score
                  1771
                  0.0e + 00
E value
                  339
Match length
                  100
% identity
                  protein kinase 1 - soybean >gi_169991 (L01453) protein
NCBI Description
                  kinase [Glycine max]
                  1072
Seq. No.
                  564 4.R1040
Contig ID
                  hrw701060550.h1
5'-most EST
                  BLASTN
Method
NCBI GI
                  q310579
                  498
BLAST score
                  0.0e + 00
E value
                  1118
Match length
                  95
% identity
NCBI Description Glycine max protein kinase 2 (SPK-2) mRNA, complete cds
                  1073
Seq. No.
                  564 5.R1040
Contig ID
5'-most EST
                  ssr700560617.hl
Method
                  BLASTN
NCBI GI
                  g310579
BLAST score
                  230
                  1.0e-126
E value
                  242
Match length
                  99
% identity
NCBI Description Glycine max protein kinase 2 (SPK-2) mRNA, complete cds
                   1074
Seq. No.
                   564 7.R1040
Contig ID
5'-most EST
                  LIB3093-053-Q1-K1-G5
                  BLASTN
Method
NCBI GI
                   q169990
BLAST score
                   154
E value
                   5.0e-81
Match length
                   229
% identity
                   98
NCBI Description Soybean protein kinase (SPK-1) mRNA, complete cds
                   1075
Seq. No.
                   567 1.R1040
Contig ID
5'-most EST
                   LIB3051-090-Q1-K1-A2
Method
                   BLASTN
NCBI GI
                   q170009
                   1494
BLAST score
E value
                   0.0e + 00
Match length
                   1727
                   99
% identity
                  Soybean late embryogenesis abundant (LEA) protein mRNA,
NCBI Description
                   complete cds
```

Seq. No.

1076

Contig ID 567_2.R1040

5'-most EST LIB3029-001-Q1-B1-F8



```
Method
                   BLASTN
NCBI GI
                   q170009
BLAST score
                   369
E value
                   0.0e + 00
Match length
                   799
                   61
% identity
                  Soybean late embryogenesis abundant (LEA) protein mRNA,
NCBI Description
                   complete cds
                   1077
Seq. No.
                   567 4.R1040
Contig ID
5'-most EST
                  LIB3065-004-Q1-N1-F2
Method
                   BLASTN
NCBI GI
                   a170009
BLAST score
                   86
E value
                   1.0e-40
Match length
                   342
% identity
                   61
                  Soybean late embryogenesis abundant (LEA) protein mRNA,
NCBI Description
                   complete cds
Seq. No.
                   1078
Contig ID
                   567 5.R1040
5'-most EST
                   LIB3029-006-Q1-B1-G3
Method
                   BLASTN
NCBI GI
                   q170009
BLAST score
                   127
E value
                   5.0e-65
Match length
                   359
% identity
                   84
                  Soybean late embryogenesis abundant (LEA) protein mRNA,
NCBI Description
                   complete cds
Seq. No.
                   1079
                   567 7.R1040
Contig ID
                   crh700850265.h1
5'-most EST
Method
                   BLASTN
                   g170009
NCBI GI
BLAST score
                   144
                   5.0e-75
E value
                   495
Match length
                   62
% identity
NCBI Description
                   Soybean late embryogenesis abundant (LEA) protein mRNA,
                   complete cds
Seq. No.
                   1080
                   567 8.R1040
Contig ID
5'-most EST
                   awf700839939.h1
Method
                   BLASTN
NCBI GI
                   g170009
BLAST score
                   123
```

Method BLASTN
NCBI GI g170009
BLAST score 123
E value 9.0e-63
Match length 301
% identity 92

NCBI Description Soybean late embryogenesis abundant (LEA) protein mRNA,

complete cds



```
Seq. No.
                   571 1.R1040
Contig ID
5'-most EST
                   uC-gmropic057g04b1
Method
                  BLASTX
NCBI GI
                   g1362046
BLAST score
                   1888
                   0.0e+00
E value
Match length
                   426
                   99
% identity
NCBI Description
                  aspartate transaminase (EC 2.6.1.1) precursor - soybean
                  >gi 710596 (L40579) mitochondrial aspartate
                   aminotransferase [Glycine max]
                  1082
Seq. No.
                   576 1.R1040
Contig ID
5'-most EST
                   g857396 FL
Method
                  BLASTX
NCBI GI
                  g857397
BLAST score
                   2283
                  0.0e + 00
E value
Match length
                   484
% identity
                   92
                  (D50870) mitotic cyclin a2-type [Glycine max]
NCBI Description
Seq. No.
                  1083
Contig ID
                  576 2.R1040
                  ncj700988646.h1
5'-most EST
Method
                  BLASTN
                  g857396
NCBI GI
BLAST score
                  124
                  2.0e-63
E value
Match length
                  220
                  90
% identity
NCBI Description Glycine max mRNA for mitotic cyclin a2-type, complete cds
Seq. No.
                  1084
Contig ID
                  577 1.R1040
5'-most EST
                  LIB3049-047-Q1-E1-A4
Method
                  BLASTX
NCBI GI
                  q857399
BLAST score
                  2091
                  0.0e+00
E value
Match length
                  427
% identity
                  96
NCBI Description
                  (D50871) mitotic cyclin b1-type [Glycine max]
Seq. No.
                  1085
Contig ID
```

577 2.R1040

5'-most EST

jC-gmf102220070f12a1

Method BLASTN NCBI GI q857398 BLAST score 306 E value 1.0e-171 Match length 354

% identity 97

NCBI Description Glycine max mRNA for mitotic cyclin b1-type, complete cds



```
Seq. No.
                   578 1.R1040
Contig ID
5'-most EST
                   LIB3056-001-Q1-B1-H5
Method
                   BLASTN
NCBI GI
                   g347458
BLAST score
                   1016
E value
                   0.0e+00
Match length
                   1048
% identity
                   99
NCBI Description
                  Glycine max brassinosteroid-regulated protein mRNA,
                   complete cds
Seq. No.
                   1087
                   578 2.R1040
Contig ID
5'-most EST
                   fua701043403.h1
Method
                   BLASTN
NCBI GI
                   g347458
BLAST score
                   361
E value
                   0.0e+00
Match length
                   712
% identity
                   90
NCBI Description
                  Glycine max brassinosteroid-regulated protein mRNA,
                  complete cds
                   1088
Seq. No.
                   578 6.R1040
Contig ID
5'-most EST
                  LIB3139-015-P1-N1-H4
Method
                  BLASTN
NCBI GI
                  g347458
BLAST score
                   216
E value
                   1.0e-118
Match length
                   365
% identity
                   94
NCBI Description
                  Glycine max brassinosteroid-regulated protein mRNA,
                  complete cds
Seq. No.
                  1089
Contig ID
                  579 1.R1040
5'-most EST
                  txt700731909.h1
Method
                  BLASTX
NCBI GI
                  g1345976
BLAST score
                  2058
                  0.0e+00
E value
Match length
                  387
% identity
                  97
NCBI Description
                  OMEGA-6 FATTY ACID DESATURASE, ENDOPLASMIC RETICULUM
                  ISOZYME 1 >gi 904152 (L43920) microsomal omega-6 desaturase
                   [Glycine max]
```

Seq. No. 1090

Contig ID 579 2.R1040 5'-most EST zpv700757037.h1

Method BLASTN NCBI GI q904151 BLAST score 375 0.0e+00E value Match length 560



% identity 94

NCBI Description Glycine max FAD2-1 microsomal omega-6 desaturase mRNA,

complete cds

Seq. No. 1091

Contig ID 579 4.R1040

5'-most EST uC-gmropic066c08b1

Seq. No. 1092

Contig ID 579_5.R1040 5'-most EST vwf700675911.h1

Method BLASTN
NCBI GI g904151
BLAST score 151
E value 3.0e-79
Match length 229
% identity 100

NCBI Description Glycine max FAD2-1 microsomal omega-6 desaturase mRNA,

complete cds

Seq. No. 1093

Contig ID 579_7.R1040 5'-most EST vwf700678878.h1

Method BLASTN
NCBI GI g904151
BLAST score 86
E value 9.0e-41
Match length 236
% identity 94

NCBI Description Glycine max FAD2-1 microsomal omega-6 desaturase mRNA,

complete cds

Seq. No. 1094

Contig ID 582 1.R1040 5'-most EST gsv701045761.h1

Method BLASTX
NCBI GI g1906002
BLAST score 2837
E value 0.0e+00
Match length 612
% identity 91

NCBI Description (U90429) nitrite reductase [Glycine max]

Seq. No. 1095

Contig ID 582 2.R1040

5'-most EST LIB3138-051-Q1-N1-F12

Method BLASTN
NCBI GI g1906001
BLAST score 276
E value 1.0e-154
Match length 518
% identity 93

NCBI Description Glycine max nitrite reductase gene, complete cds

Seq. No. 1096

Contig ID 582_3.R1040

5'-most EST jC-gmle01810077g08d1



```
Method
                  BLASTN
NCBI GI
                  g1906001
                  189
BLAST score
                  1.0e-102
E value
Match length
                  189
                  100
% identity
NCBI Description Glycine max nitrite reductase gene, complete cds
                  1097
Seq. No.
                  582 5.R1040
Contig ID
5'-most EST
                  ssr700556787.h1
```

BLASTN Method g1906001 NCBI GI BLAST score 285 1.0e-159 E value Match length 373

96 % identity

NCBI Description Glycine max nitrite reductase gene, complete cds

1098 Seq. No.

585 1.R1040 Contig ID

5'-most EST LIB3049-002-Q1-E1-B9

Method BLASTN NCBI GI q310575 180 BLAST score 3.0e-96 E value Match length 844 % identity 88

NCBI Description Glycine max nodulin-26 mRNA, complete cds

1099 Seq. No.

Contig ID 585 2.R1040

5'-most EST LIB3039-023-Q1-E1-A3

Method BLASTN NCBI GI g310575 BLAST score 209 E value 1.0e-113 Match length 1016 % identity 90

NCBI Description Glycine max nodulin-26 mRNA, complete cds

1100 Seq. No.

587 1.R1040 Contig ID leu701153015.h1 5'-most EST

BLASTN Method NCBI GI g310577 BLAST score 918 0.0e+00E value 963 Match length 99 % identity

NCBI Description Glycine max nodulin-26 mRNA, complete cds

Seq. No. 1101

588 1.R1040 Contig ID 5'-most EST g409368 FL Method BLASTN NCBI GI g409368

die.



BLAST score 715 E value 0.0e+00 Match length 867 % identity 98

NCBI Description Glycine max mutant nitrate reductase mRNA, 3' end

Seq. No. 1102

Contig ID 588 2.R1040

5'-most EST LIB3138-056-Q1-N1-G10

Method BLASTX
NCBI GI g730142
BLAST score 657
E value 5.0e-69
Match length 141
% identity 91

NCBI Description NITRATE REDUCTASE 2 (NR-2) >gi_392992 (U01029) nitrate

reductase [Phaseolus vulgaris]

Seq. No. 1103

Contig ID 588_3.R1040 5'-most EST smc700749263.h1

Method BLASTN
NCBI GI g392991
BLAST score 124
E value 5.0e-63
Match length 396
% identity 83

NCBI Description Phaseolus vulgaris nitrate reductase (PVNR2) gene, complete

cds

Seq. No. 1104

Contig ID 589_1.R1040

5'-most EST LIB3109-035-Q1-K1-B8

Method BLASTX
NCBI GI g1345979
BLAST score 2322
E value 0.0e+00
Match length 424
% identity 99

NCBI Description OMEGA-6 FATTY ACID DESATURASE, CHLOROPLAST PRECURSOR

>gi 459962 (L29215) plastid omega-6 desaturase [Glycine

max]

Seq. No. 1105

Contig ID 589 5.R1040

5'-most EST jC-gmle01810089d01a1

Method BLASTN
NCBI GI g459961
BLAST score 259
E value 1.0e-144
Match length 352
% identity 98

NCBI Description Glycine max proplastid omega-6 desaturase mRNA, complete

cds

Seq. No. 1106

Contig ID 592 1.R1040

```
5'-most EST
                   LIB3109-037-Q1-K1-D4
Method
                   BLASTN
NCBI GI
                   g170041
BLAST score
                   920
E value
                   0.0e+00
Match length
                   1093
% identity
                   97
NCBI Description Glycine max cv Prize protein kinase mRNA
Seq. No.
                   1107
Contig ID
                   592 2.R1040
5'-most EST
                   LIB3109-013-Q1-K1-A1
Method
                   BLASTN
NCBI GI
                   g170042
BLAST score
                   1216
E value
                   0.0e+00
Match length
                   1216
% identity
                   100
NCBI Description Glycine max cv Prize protein kinase mRNA
Seq. No.
                   1108
Contig ID
                   592 3.R1040
5'-most EST
                  uaw700665730.h1
Method
                  BLASTN
NCBI GI
                  g170041
BLAST score
                   255
E value
                  1.0e-141
Match length
                  259
                  100
% identity
NCBI Description Glycine max cv Prize protein kinase mRNA
Seq. No.
                  1109
                  592 4.R1040
Contig ID
5'-most EST
                  leu701155312.h1
Method
                  BLASTN
NCBI GI
                  g170042
BLAST score
                  246
E value
                  1.0e-136
Match length
                  341
                  96
% identity
NCBI Description Glycine max cv Prize protein kinase mRNA
Seq. No.
                  1110
Contig ID
                  594 1.R1040
5'-most EST
                  g736338 FL
Method
                  BLASTX
```

1

NCBI GI q1171965 BLAST score 4069 E value 0.0e+00

Match length 814 % identity 97

NCBI Description PHOSPHATIDYLINOSITOL 3-KINASE, ROOT ISOFORM (PI3-KINASE) (PTDINS-3-KINASE) (PI3K) (SPI3K-5) >gi 736339 (L27265)

phosphatidylinositol 3-kinase [Glycine max]

Seq. No. 1111

Contig ID 596 1.R1040



```
5'-most EST g515750_FL
Method BLASTX
NCBI GI g1172495
BLAST score 3387
E value 0.0e+00
Match length 1106
% identity 94
NCBI Description BLASTOCHEOM
```

NCBI Description PHYTOCHROME A >gi_515749 (L34842) phytochrome A [Glycine max] >gi_515751 (L34844) phytochrome A [Glycine max]

Seq. No. 1112

Contig ID 596_2.R1040

5'-most EST fC-gmse700654835a6

Method BLASTN
NCBI GI g515748
BLAST score 219
E value 1.0e-120
Match length 435
% identity 94

NCBI Description Soybean chloroplast phytochrome A (phyA) gene, complete cds

Seq. No. 1113

Contig ID 596 3.R1040

5'-most EST uC-gmflminsoy012f11b1

Method BLASTX
NCBI GI g3953465
BLAST score 452
E value 1.0e-44
Match length 160
% identity 55

NCBI Description (AC002328) F20N2.10 [Arabidopsis thaliana]

Seq. No. 1114

Contig ID 596 4.R1040

5'-most EST fC-gmst700660709a4

Method BLASTN
NCBI GI g515750
BLAST score 320
E value 1.0e-180
Match length 719
% identity 91

NCBI Description Soybean phytochrome A (phyA) mRNA, complete cds

Seq. No. 1115

597 1.R1040 Contig ID 5'-most EST g170043 FL Method BLASTX NCBI GI q117519 BLAST score 2972 0.0e+00E value Match length 570 % identity 100

NCBI Description PHYTOENE DEHYDROGENASE PRECURSOR (PHYTOENE DESATURASE) >gi_99945_pir__A39597 phytoene dehydrogenase (EC 1.3.-.-) -

soybean >gi_170044 (M64704) phytoene desaturase [Glycine

maxl



```
Seq. No.
                   1116
Contig ID
                   597 2.R1040
5'-most EST
                   fua701041583.h1
Method
                   BLASTN
NCBI GI
                   q170043
BLAST score
                   143
E value
                   9.0e-75
Match length
                   219
% identity
                   94
NCBI Description Soybean phytoene desaturase mRNA, complete cds
Seq. No.
                   1117
Contig ID
                   599 1.R1040
5'-most EST
                   ssr700560806.h1
Method
                   BLASTX
NCBI GI
                   q478809
BLAST score
                   2356
E value
                   0.0e+00
Match length
                   462
% identity
                   97
NCBI Description
                  protein kinase 6 (EC 2.7.1.-) - soybean >gi 170047 (M67449)
                   protein kinase [Glycine max] >gi 444789 prf 1908223A
                  protein kinase [Glycine max]
Seq. No.
                   1118
Contig ID
                   599 2.R1040
5'-most EST
                   LIB3093-030-Q1-K1-E1
Method
                   BLASTN
NCBI GI
                   g170046
BLAST score
                   177
E value
                   1.0e-94
Match length
                   345
% identity
                   92
NCBI Description Glycine max protein kinase (PK6) mRNA, complete cds
Seq. No.
                  1119
                   600 1.R1040
Contig ID
5'-most EST
                  hrw701060433.h1
Method
                  BLASTX
NCBI GI
                  g2497538
BLAST score
                  2531
E value
                  0.0e+00
Match length
                  511
% identity
                  98
NCBI Description
                  PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi 466350 (L08632)
                  pyruvate kinase [Glycine max]
Seq. No.
                  1120
                  600 3.R1040
Contig ID
5'-most EST
                  uC-gmropic054f07b1
Method
                  BLASTN
                  q170050
```

Method BLASTN
NCBI GI g170050
BLAST score 314
E value 1.0e-176
Match length 318
% identity 100

NCBI Description Soybean pyruvate kinase mRNA, complete cds



Seq. No. 1121

Contig ID 600 5.R1040

5'-most EST uC-gmropic047d08b1

Method BLASTN
NCBI GI g170050
BLAST score 301
E value 1.0e-169
Match length 434

% identity 94

NCBI Description Soybean pyruvate kinase mRNA, complete cds

Seq. No. 1122

Contig ID 600 8.R1040 5'-most EST ncj700975678.h1

Method BLASTN
NCBI GI g170050
BLAST score 147
E value 5.0e-77
Match length 233
% identity 98

NCBI Description Soybean pyruvate kinase mRNA, complete cds

Seq. No. 1123

Contig ID 601 1.R1040

5'-most EST LIB3109-036-Q1-K1-C10

Method BLASTN
NCBI GI g414831
BLAST score 776
E value 0.0e+00
Match length 815
% identity 99

NCBI Description Glycine max (Rablp) mRNA, complete cds

Seq. No. 1124

602 1.R1040 Contig ID 5'-most EST g414833 FL Method BLASTN NCBI GI g414833 BLAST score 1076 E value 0.0e + 00Match length 1118 100 % identity

NCBI Description Glycine max (Rab7p) mRNA, complete cds

Seq. No. 1125

Contig ID 602_2.R1040 5'-most EST fua701039675.h1 Method BLASTN

Method BLASTN
NCBI GI 9414833
BLAST score 195
E value 1.0e-106
Match length 227
% identity 96

NCBI Description Glycine max (Rab7p) mRNA, complete cds

Seq. No. 1126



Contig ID 603 1.R1040 awf700838208.h1 5'-most EST Method BLASTN NCBI GI a170051 BLAST score 851 E value 0.0e+00Match length 884 % identity 99 NCBI Description Glycine max RNA polymerase II fifth largest subunit mRNA, complete cds 25. Seq. No. 1127 Contig ID 603 2.R1040 5'-most EST uC-gmropic009g08b1 BLASTN Method NCBI GI g170051 BLAST score 354 E value 0.0e+00 Match length 433 % identity 96 NCBI Description Glycine max RNA polymerase II fifth largest subunit mRNA, complete cds Seq. No. 1128 Contig ID 604 1.R1040 5'-most EST LIB3107-038-Q1-K1-E12 Method BLASTN NCBI GI g170053 BLAST score 545 E value 0.0e + 00Match length 682 % identity 95 NCBI Description Soybean ribosomal protein S11 mRNA, 3' end Seq. No. 1129 604 2.R1040 Contig ID 5'-most EST LIB3039-047-Q1-E1-E3 Method BLASTN NCBI GI q170053 BLAST score 224 E value 1.0e-122 Match length 416

% identity 88

NCBI Description Soybean ribosomal protein S11 mRNA, 3' end

Seq. No. 1130

Contig ID 604 3.R1040

5'-most EST LIB3049-056-Q1-E1-F10

Method BLASTN NCBI GI q170053 BLAST score 432 E value 0.0e+00Match length 579 94 % identity

NCBI Description Soybean ribosomal protein S11 mRNA, 3' end

Seq. No. 1131



```
Contig ID
                   604 5.R1040
                   LIB3039-015-Q1-E1-H4
5'-most EST
Method
                   BLASTN
NCBI GI
                   q170053
BLAST score
                   175
E value
                   1.0e-93
Match length
                   211
% identity
                   96
NCBI Description Soybean ribosomal protein S11 mRNA, 3' end
Seq. No.
                   1132
Contig ID
                   604 9.R1040
                   LIB3040-029-Q1-E1-H5
5'-most EST
Method
                   BLASTN
                   g170053
NCBI GI
BLAST score
                   91
E value
                   1.0e-43
Match length
                   195
% identity
                   87
NCBI Description Soybean ribosomal protein S11 mRNA, 3' end
Seq. No.
                   1133
Contig ID
                   605 1.R1040
5'-most EST
                   ssr700553763.h1
Method
                  BLASTN
NCBI GI
                  q454178
BLAST score
                   1264
E value
                   0.0e + 00
Match length
                   1285
% identity
                   99
NCBI Description Glycine max rubisco-associated protein mRNA, complete cds
Seq. No.
                  1134
                  606 1.R1040
Contig ID
5'-most EST
                  LIB3040-030-Q1-E1-B7
Method
                  BLASTX
NCBI GI
                  g3915032
BLAST score
                  1814
E value
                  0.0e+00
Match length
                  411
% identity
                  86
NCBI Description
                  ACYL-[ACYL-CARRIER PROTEIN] DESATURASE PRECURSOR
                  (STEAROYL-ACP DESATURASE) >gi_508603 (L34346) stearoyl-acyl
                  carrier protein desaturase [Glycine max]
Seq. No.
                  1135
Contig ID
                  606 2.R1040
5'-most EST
                  uC-gmflminsoy028f03b1
                  BLASTN
Method
NCBI GI
                  q508602
BLAST score
                  268
```

E value 1.0e-149 Match length 348 % identity 94

NCBI Description Glycine max stearoyl-acyl carrier protein desaturase

(SACPD) mRNA, complete cds



```
Seq. No.
                   1136
Contig ID
                   607 1.R1040
                   uxk700671927.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q530207
BLAST score
                   4384
E value
                   0.0e + 00
Match length
                   896
% identity
                   98
NCBI Description
                  (L35272) heat shock protein [Glycine max]
Seq. No.
                   1137
Contig ID
                   608 1.R1040
5'-most EST
                   zpv700757989.h1
Method
                   BLASTX
NCBI GI
                   g548900
BLAST score
                   1973
E value
                   0.0e + 00
Match length
                   504
% identity
                   78
NCBI Description
                  SUCROSE-BINDING PROTEIN PRECURSOR (SBP)
                  >gi_322691 pir JQ1730 62K sucrose-binding protein
                  precursor - soybean >gi 170064 (L06038) glucose binding
                  protein [Glycine max]
Seq. No.
                  1138
Contig ID
                   608 2.R1040
5'-most EST
                   zpv700763038.h1
Method
                  BLASTN
NCBI GI
                  q1431744
BLAST score
                   392
E value
                  0.0e+00
Match length
                  737
% identity
                   98
NCBI Description
                  Glycine max sucrose binding protein (sbp) mRNA, complete
Seq. No.
                  1139
Contig ID
                  608 3.R1040
5'-most EST
                  uaw700662248.h1
Method
                  BLASTN
NCBI GI
                  g1431744
BLAST score
                  155
E value
                  1.0e-81
Match length
                  519
% identity
                  84
                  Glycine max sucrose binding protein (sbp) mRNA, complete
NCBI Description
Seq. No.
                  1140
```

609 1.R1040 Contig ID 5'-most EST wrq700785963.h2 Method BLASTN

NCBI GI g170073 BLAST score 433 E value 0.0e + 00Match length 829



```
% identity 92
NCBI Description Soybean calmodulin (SCaM-3) mRNA, complete cds
Seq. No. 1141
Contig ID 609_2.R1040
5'-most EST LIB3072-053-Q1-E1-E12
Method BLASTN
```

NCBI GI g170069
BLAST score 705
E value 0.0e+00
Match length 738
% identity 99

NCBI Description Soybean calmodulin (SCaM-1) mRNA, complete cds

Seq. No. 1142

Contig ID 609 3.R1040

5'-most EST LIB3106-104-Q1-K1-F4

Method BLASTN
NCBI GI g170073
BLAST score 431
E value 0.0e+00
Match length 435
% identity 100

NCBI Description Soybean calmodulin (SCaM-3) mRNA, complete cds

Seq. No. 1143

Contig ID 609 5.R1040

5'-most EST uC-gmflminsoy118b07b1

Method BLASTN
NCBI GI g170069
BLAST score 265
E value 1.0e-147
Match length 321
% identity 96

NCBI Description Soybean calmodulin (SCaM-1) mRNA, complete cds

Seq. No. 1144

Contig ID 610 1.R1040

5'-most EST LIB3049-024-Q1-E1-H5

Method BLASTN
NCBI GI g170071
BLAST score 668
E value 0.0e+00
Match length 839
% identity 97

NCBI Description Soybean calmodulin (SCaM-2) mRNA, complete cds

Seq. No. 1145

Contig ID 610_2.R1040 5'-most EST kl1701205314.h1

Method BLASTN
NCBI GI g4103958
BLAST score 460
E value 0.0e+00
Match length 788
% identity 92

NCBI Description Phaseolus vulgaris calmodulin (CaM) mRNA, PvCaM-2 allele,

Match length

% identity

901

87



complete cds

```
Seq. No.
                   1146
 Contig ID
                   612 1.R1040
 5'-most EST
                   g170075 FL
Method
                   BLASTN
NCBI GI
                   g170075
BLAST score
                   764
E value
                   0.0e+00
Match length
                   1352
% identity
                   100
NCBI Description Soybean calmodulin (SCaM-4) mRNA, complete cds
Seq. No.
                   1147
Contig ID
                   612 3.R1040
5'-most EST
                   LIB3053-012-Q1-N1-F1
Method
                   BLASTN
NCBI GI
                   g170075
BLAST score
                   105
E value
                   6.0e-52
Match length
                   201
% identity
                   100
NCBI Description Soybean calmodulin (SCaM-4) mRNA, complete cds
Seq. No.
                   1148
Contig ID
                   613_1.R1040
5'-most EST
                   g532289 FL
Method
                   BLASTX
NCBI GI
                   g2492792
BLAST score
                   2806
E value
                   0.0e+00
Match length
                   549
% identity
                   100
NCBI Description
                  LATE NODULIN 56 (N-56) >gi_532290_dbj_BAA07212_ (D38015)
                   soybean late nodulin [Glycine max]
Seq. No.
                   1149
Contig ID
                   614 1.R1040
5'-most EST
                   LIB3049-004-Q1-E1-G9
Method
                   BLASTN
NCBI GI
                   g1161251
BLAST score
                   852
E value
                   0.0e+00
Match length
                   1364
% identity
NCBI Description
                  Glycine max nucleosome assembly protein 1 (SNAP-1) mRNA,
                  complete cds
Seq. No.
                  1150
                  614 2.R1040
Contig ID
5'-most EST
                  leu701157276.h1
Method
                  BLASTN
NCBI GI
                  g1161251
BLAST score
                  291
E value
                  1.0e-162
```



NCBI Description Glycine max nucleosome assembly protein 1 (SNAP-1) mRNA, complete cds

Seq. No. 1151

Contig ID 614 4.R1040

5'-most EST uC-gmflminsoy043h04b1

Method BLASTN
NCBI GI g1161251
BLAST score 120
E value 6.0e-61
Match length 219
% identity 94

NCBI Description Glycine max nucleosome assembly protein 1 (SNAP-1) mRNA,

complete cds

Seq. No. 1152

Contig ID 614_7.R1040 5'-most EST crh700850781.h1

Method BLASTN
NCBI GI g1161251
BLAST score 130
E value 6.0e-67
Match length 200
% identity 93

NCBI Description Glycine max nucleosome assembly protein 1 (SNAP-1) mRNA,

complete cds

Seq. No. 1153

Contig ID 617_1.R1040

5'-most EST uC-gmrominsoy203b07b1

Method BLASTN
NCBI GI g1220521
BLAST score 953
E value 0.0e+00
Match length 1070
% identity 96

NCBI Description Glycine max TATA-box binding protein (STBP1) mRNA, complete

cds

Seq. No. 1154

Contig ID 617_2.R1040 5'-most EST zhf700952839.h1

Method BLASTN
NCBI GI g1220521
BLAST score 123
E value 1.0e-62
Match length 223
% identity 89

NCBI Description Glycine max TATA-box binding protein (STBP1) mRNA, complete

cds

Seq. No. 1155

Contig ID 617_3.R1040

5'-most EST LIB3087-005-Q1-K1-B7

Method BLASTN
NCBI GI g1220521
BLAST score 121

289



E value 2.0e-61 Match length 228 % identity 96

NCBI Description Glycine max TATA-box binding protein (STBP1) mRNA, complete

cds

Seq. No. 1156

Contig ID 617_4.R1040 5'-most EST uxk700669017.h1

Seq. No. 1157

Contig ID 617_5.R1040 5'-most EST ssr700560377.h1

Method BLASTN
NCBI GI g1220521
BLAST score 131
E value 1.0e-67
Match length 240
% identity 100

NCBI Description Glycine max TATA-box binding protein (STBP1) mRNA, complete

cds

Seq. No. 1158

Contig ID 618_1.R1040 5'-most EST wvk700685392.h1

Method BLASTX
NCBI GI g986969
BLAST score 1765
E value 0.0e+00
Match length 362
% identity 96

NCBI Description (L28005) TGACG-motif-binding protein [Glycine max]

Seq. No. 1159

Contig ID 618_3.R1040 5'-most EST ary700764344.h1

Method BLASTN
NCBI GI g986968
BLAST score 167
E value 6.0e-89
Match length 296
% identity 94

NCBI Description Glycine max TGACG-motif binding protein (STGA1) mRNA,

complete cds

Seq. No. 1160

Contig ID 622_1.R1040

5'-most EST LIB3028-049-Q1-B1-F5

Method BLASTN
NCBI GI g4100432
BLAST score 258
E value 1.0e-142
Match length 524
% identity 97

NCBI Description Glycine max beta-glucosidase mRNA, partial cds

Seq. No. 1161



Contig ID 624_1.R1040 5'-most EST awf700838867.h1

Method BLASTN
NCBI GI g2270987
BLAST score 798
E value 0.0e+00
Match length 798
% identity 100

NCBI Description Glycine max Em protein mRNA, complete cds

Seq. No. 1162

Contig ID 624 2.R1040 5'-most EST awf700839943.h1

Method BLASTN
NCBI GI g2270987
BLAST score 228
E value 1.0e-125
Match length 394
% identity 92

NCBI Description Glycine max Em protein mRNA, complete cds

Seq. No. 1163

Contig ID 625_1.R1040
5'-most EST g4102689_FL
Method BLASTN
NCBI GI g4102689
BLAST score 885
E value 0.0e+00
Match length 885

% identity 100
NCBI Description Glycine max 24 kDa seed maturation protein mRNA, complete

Seq. No. 1164

Contig ID 626_1.R1040

5'-most EST LIB3094-100-Q1-K1-F2

Method BLASTN
NCBI GI g2270989
BLAST score 729
E value 0.0e+00
Match length 729
% identity 100

NCBI Description Glycine max dehydrin (GmPM12) mRNA, complete cds

Seq. No.

1165

Contig ID 626_2.R1040

5'-most EST LIB3051-006-Q1-K1-F3

Method BLASTN
NCBI GI g2270989
BLAST score 293
E value 1.0e-164
Match length 384
% identity 99

NCBI Description Glycine max dehydrin (GmPM12) mRNA, complete cds

Seq. No. 1

1166

Contig ID 627 1.R1040



```
LIB3072-017-Q1-E1-G4
5'-most EST
                  BLASTN
Method
                  g2306978
NCBI GI
BLAST score
                  490
                  0.0e + 00
E value
                  502
Match length
                  99
% identity
                 Glycine max metallothionein-II protein (PGMPM19) mRNA,
NCBI Description
                  complete cds
                  1167
Seq. No.
                  627 2.R1040
Contig ID
                  LIB3072-028-Q1-E1-G2
5'-most EST
                  BLASTN
Method
NCBI GI
                  g2270991
BLAST score
                   287
                   1.0e-160
E value
                   386
Match length
                   100
% identity
NCBI Description Glycine max metallothionein-II protein mRNA, complete cds
                   1168
Seq. No.
                   627 3.R1040
Contiq ID
                   LIB3072-059-Q1-K1-G7
5'-most EST
                   BLASTN
Method
                   q2270991
NCBI GI
                   413
BLAST score
                   0.0e + 00
E value
                   413
Match length
                   100
% identity
NCBI Description Glycine max metallothionein-II protein mRNA, complete cds
                   1169
Seq. No.
                   628 1.R1040
Contig ID
                   uxk700668563.h1
5'-most EST
                   BLASTN
Method
                   g2270993
NCBI GI
                   819
BLAST score
                   0.0e + 00
E value
                   964
Match length
                   99
% identity
                   Glycine max Ca+2-binding EF hand protein (GmPM13) mRNA,
NCBI Description
                   complete cds
                   1170
Seq. No.
                    629 1.R1040
Contig ID
                   awf700843103.h1
 5'-most EST
                   BLASTN
Method
                    q4102691
NCBI GI
                    633
 BLAST score
 E value
                    0.0e + 00
                    633
Match length
 % identity
                   Glycine max late-embryogenesis abundant protein mRNA,
 NCBI Description
```

Seq. No.

292

complete cds

1171



```
629 2.R1040
Contig ID
5'-most EST
                  crh700854653.hl
                  BLASTN
Method
                  q4102691
NCBI GI
BLAST score
                  275
                  1.0e-153
E value
                  532
Match length
                  91
% identity
NCBI Description Glycine max late-embryogenesis abundant protein mRNA,
                  complete cds
                   1172
Seq. No.
                   629 3.R1040
Contig ID
                  LIB3170-004-Q1-K1-E10
5'-most EST
                  BLASTN
Method
NCBI GI
                   g4102691
BLAST score
                   240
                   1.0e-132
E value
                   519
Match length
                   93
% identity
NCBI Description Glycine max late-embryogenesis abundant protein mRNA,
                   complete cds
                   1173
Seq. No.
                   630 1.R1040
Contig ID
5'-most EST
                   zpv700757378.h1
Method
                   BLASTN
                   g4097895
NCBI GI
BLAST score
                   728
                   0.0e+00
E value
                   728
Match length
                   100
% identity
NCBI Description Glycine max napin-type 2S albumin 3 mRNA, complete cds
                   1174
Seq. No.
                   630_2.R1040
Contig ID
                   uC-gmropic066f10b1
5'-most EST
                   BLASTX
Method
                   q4063749
NCBI GI
                   239
BLAST score
                   6.0e-20
E value
                   92
Match length
                   27
% identity
                   (AC005851) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   1175
Seq. No.
                    630 4.R1040
Contig ID
                   LIB3073-008-Q1-K1-G3
 5'-most EST
                   BLASTN
Method
NCBI GI
                    q4097895
 BLAST score
                    129
                    8.0e-66
 E value
                    241
Match length
                    98
 % identity
```

Seq. No. 1176

NCBI Description Glycine max napin-type 2S albumin 3 mRNA, complete cds



```
630 6.R1040
Contig ID
5'-most EST
                  vwf700678862.h1
Method
                  BLASTN
NCBI GI
                  q2305019
BLAST score
                  172
                  6.0e-92
E value
Match length
                  176
                  99
% identity
NCBI Description Glycine max 2S albumin pre-propeptide mRNA, complete cds
                  1177
Seq. No.
Contig ID
                  631 1.R1040
5'-most EST
                  LIB3051-110-Q1-K1-F5
Method
                  BLASTX
NCBI GI
                  g2253384
```

BLAST score 2739 0.0e+00 E value Match length 539 % identity 99

(AF007100) biotin carboxylase precursor [Glycine max] NCBI Description

Contig ID 631 2.R1040 5'-most EST fua701040021.h1 Method BLASTX NCBI GI q134892 BLAST score 654 E value 2.0e-68 Match length 194 % identity 68

Seq. No.

SIGNAL RECOGNITION PARTICLE RECEPTOR ALPHA SUBUNIT NCBI Description

(SR-ALPHA) (DOCKING PROTEIN ALPHA) (DP-ALPHA)

>gi_88607_pir__A29440 signal recognition particle receptor - human >gi_30866 emb_CAA29608 (X06272) docking protein [Homo sapiens] >gi_4507223_ref_NP_003130.1_pSRPR_signal

recognition particle receptor ('docking protein')

Seq. No. 1179 631 3.R1040 Contig ID g3219360 FL 5'-most EST Method BLASTX NCBI GI g3219361 BLAST score 2741 E value 0.0e + 00Match length 539 99 % identity

NCBI Description (AF068249) biotin carboxylase precursor [Glycine max]

Seq. No. 1180 631 6.R1040 Contig ID 5'-most EST LIB3055-001-Q1-B1-F6 Method BLASTN NCBI GI g3219360 BLAST score 182 E value 5.0e-98

Match length 262 % identity 93



NCBI Description Glycine max biotin carboxylase precursor (accC-3) mRNA, nuclear gene encoding chloroplast protein, complete cds

Seq. No. 1181

Contig ID 632 1.R1040

5'-most EST uC-gmropic022d10b1

Method BLASTX
NCBI GI g2245683
BLAST score 1680
E value 0.0e+00
Match length 339
% identity 99

NCBI Description (AF007211) peroxidase precursor [Glycine max]

Seq. No. 1182

Contig ID 632_2.R1040 5'-most EST trc700563017.h1

Method BLASTN
NCBI GI g2245682
BLAST score 289
E value 1.0e-161
Match length 341
% identity 96

NCBI Description Glycine max peroxidase precursor (GMIPER1) mRNA, complete

cds

Seq. No. 1183

Contig ID 633 1.R1040 5'-most EST g4102189 FL Method BLASTN g4102189 NCBI GI BLAST score 1208 E value 0.0e+00Match length 1220 % identity 98

NCBI Description Glycine max 35 kDa seed maturation protein (pGmPM18) mRNA,

complete cds

Seq. No. 1184

Contig ID 633_2.R1040 5'-most EST 6HA-02-Q1-B1-G6

Method BLASTN
NCBI GI g4102189
BLAST score 267
E value 1.0e-148
Match length 461
% identity 86

NCBI Description Glycine max 35 kDa seed maturation protein (pGmPM18) mRNA,

complete cds

Seq. No. 1185

Contig ID 635_1.R1040

5'-most EST LIB3167-042-P1-K1-G11

Method BLASTN
NCBI GI g2738559
BLAST score 715
E value 0.0e+00



Match length 730 % identity 100

NCBI Description Glycine max ribosomal protein L22 mRNA, nuclear gene

encoding chloroplast protein, partial cds

Seq. No. 1186

Contig ID 635_2.R1040 5'-most EST epx701105174.h1

Method BLASTN
NCBI GI g2738559
BLAST score 264
E value 1.0e-146
Match length 347
% identity 94

NCBI Description Glycine max ribosomal protein L22 mRNA, nuclear gene

encoding chloroplast protein, partial cds

Seq. No. 1187

Contig ID 636 1.R1040 5'-most EST zhf700952296.h1

Method BLASTN
NCBI GI g2388688
BLAST score 958
E value 0.0e+00
Match length 1105
% identity 97

NCBI Description Glycine max GH1 protein (GH1) mRNA, partial cds

Seq. No. 1188

Contig ID 636_2.R1040

5'-most EST LIB3107-063-Q1-K1-B2

Method BLASTN
NCBI GI g2388688
BLAST score 308
E value 1.0e-172
Match length 825
% identity 90

NCBI Description Glycine max GH1 protein (GH1) mRNA, partial cds

Seq. No. 1189

Contig ID 636_3.R1040

5'-most EST LIB3109-031-Q1-K1-E12

Method BLASTN
NCBI GI g2388688
BLAST score 214
E value 1.0e-116
Match length 385
% identity 89

NCBI Description Glycine max GH1 protein (GH1) mRNA, partial cds

Seq. No. 1190

Contig ID 636 4.R1040 5'-most EST fua701040689.h1

Method BLASTN
NCBI GI g2388688
BLAST score 71
E value 2.0e-31



Match length 143 % identity 87

NCBI Description Glycine max GH1 protein (GH1) mRNA, partial cds

Seq. No.

1191

Contig ID

636 6.R1040

5'-most EST

jC-gmro02910006g06d1

Seq. No.

1192

Contig ID

636 9.R1040

5'-most EST

LIB3039-041-Q1-E1-H8

Seq. No.

1193

Contig ID 5'-most EST Method

637 1.R1040 g2565258_FL BLASTX

NCBI GI BLAST score E value

g2565259 1641 0.0e + 00

Match length % identity

329

NCBI Description

(AF019115) putative cadmium-transporting ATPase [Glycine

max]

Seq. No.

1194

Contig ID

638 1.R1040

5'-most EST

jC-gmro02910069c06a1

Method BLASTN g3789722 NCBI GI BLAST score 1223 0.0e + 00E value Match length 1324 % identity

NCBI Description Glycine max peroxidase (PC7) mRNA, complete cds

Seq. No.

1195

Contig ID 5'-most EST 639 1.R1040 jex700903127.hl

Method NCBI GI BLAST score BLASTX g4432860 1158

E value Match length 0.0e+00637

% identity NCBI Description 32 (AC006300) putative glucose-induced repressor protein

[Arabidopsis thaliana]

Seq. No.

1196

Contig ID

639 2.R1040

5'-most EST

LIB3028-042-Q1-B1-F1

Method NCBI GI BLASTX g4432860

BLAST score E value Match length 1160 1.0e-127

368 25 % identity

NCBI Description (AC006300) putative glucose-induced repressor protein



[Arabidopsis thaliana]

```
1197
Seq. No.
                   639 3.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy026h10b1
Method
                  BLASTX
NCBI GI
                  g4432860
BLAST score
                  522
                   7.0e-53
E value
                  197
Match length
                   59
% identity
                  (AC006300) putative glucose-induced repressor protein
NCBI Description
                   [Arabidopsis thaliana]
                   1198
Seq. No.
Contig ID
                   639 4.R1040
5'-most EST
                   jC-gmst02400023g06a1
Method
                  BLASTX
NCBI GI
                   g2407790
BLAST score
                  174
                   2.0e-12
E value
Match length
                  87
% identity
                   47
NCBI Description
                  (AF019910) grr1 [Glycine max]
Seq. No.
                   1199
                   639 5.R1040
Contig ID
5'-most EST
                  LIB3170-046-Q1-K1-G11
Method
                   BLASTN
NCBI GI
                   q2407789
BLAST score
                   242
E value
                   1.0e-134
Match length
                   261
                   99
% identity
NCBI Description
                  Glycine max grr1 (grr1) mRNA, complete cds
                   1200
Seq. No.
                   641 1.R1040
Contig ID
                   g2895197 FL
5'-most EST
Method
                  BLASTX
NCBI GI
```

q3913525 BLAST score 5544 E value 0.0e + 00Match length 1088 % identity 99

NCBI Description DNA POLYMERASE DELTA CATALYTIC CHAIN >qi 2895198 (AF020193)

DNA polymerase delta [Glycine max]

Seq. No. 1201 641 3.R1040 Contig ID leu701151641.h1 5'-most EST Method BLASTN NCBI GI g2895197 BLAST score 140

E value 5.0e-73 Match length 156 97 % identity



NCBI Description Glycine max DNA polymerase delta (Pol delta) mRNA, complete cds

Seq. No. 1202

Contig ID 642_1.R1040

5'-most EST LIB3170-048-Q1-J1-E7

Method BLASTN
NCBI GI g2444419
BLAST score 1114
E value 0.0e+00
Match length 1141
% identity 99

NCBI Description Glycine max ribosome-associated protein p40 mRNA, complete

cds

Seq. No. 1203

Contig ID 642_2.R1040

5'-most EST LIB3170-036-Q1-K2-F2

Method BLASTN
NCBI GI g2815245
BLAST score 94
E value 4.0e-45
Match length 289
% identity 85

NCBI Description C.arietinum mRNA for class I type 2 metallothionein (clone:

CanMT-2)

Seq. No. 1204

Contig ID 642_3.R1040

5'-most EST jC-gmst02400076g08a1

Method BLASTN
NCBI GI g2815245
BLAST score 86
E value 2.0e-40
Match length 242
% identity 84

NCBI Description C.arietinum mRNA for class I type 2 metallothionein (clone:

CanMT-2)

Seq. No. 1205

Contig ID 642_4.R1040 5'-most EST 2hf700957326.h1

Method BLASTN
NCBI GI g2444419
BLAST score 386
E value 0.0e+00
Match length 494
% identity 95

NCBI Description Glycine max ribosome-associated protein p40 mRNA, complete

cds

Seq. No. 1206

Contig ID 642_7.R1040

5'-most EST LIB3170-019-Q1-J1-H3

Method BLASTN
NCBI GI g2444419
BLAST score 176



```
3.0e-94
E value
                  297
Match length
                  92
% identity
                 Glycine max ribosome-associated protein p40 mRNA, complete
NCBI Description
                  1207
Seq. No.
                  642 8.R1040
Contig ID
                  LIB3107-010-Q1-K1-A2
5'-most EST
                  1208
```

Seq. No. 642 9.R1040 Contig ID gsv701054621.h1 5'-most EST

1209 Seq. No. 643 1.R1040 Contig ID wvk700684093.hl 5'-most EST BLASTX Method

g2738982 NCBI GI 2518 BLAST score 0.0e + 00E value 514 Match length 95 % identity

NCBI Description (AF022157) CYP71A10 [Glycine max]

1210 Seq. No. 644 1.R1040 Contig ID g5509129 5'-most EST BLASTX Method g2738996 NCBI GI 2907 BLAST score 0.0e+00E value 561 Match length 100

% identity NCBI Description (AF022457) CYP97B2p [Glycine max]

1211 Seq. No. 645 1.R1040 Contig ID

LIB3139-049-P1-N1-H10 5'-most EST

BLASTX Method g2738998 NCBI GI 2669 BLAST score 0.0e + 00E value 509 Match length % identity

NCBI Description (AF022458) CYP98A2p [Glycine max]

1212 Seq. No.

646 1.R1040 Contig ID leu701146357.hl 5'-most EST

Method BLASTX g2739000 NCBI GI 2388 BLAST score 0.0e+00E value Match length 510 92 % identity

NCBI Description (AF022459) CYP71D10p [Glycine max]

Match length

% identity

518

NCBI Description (AF022462) CYP93Clp [Glycine max]

88



```
Seq. No.
                  1213
Contig ID
                  647 1.R1040
                  LIB3049-016-Q1-E1-A5
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2739002
BLAST score
                  2331
                  0.0e+00
E value
                  499
Match length
                  91
% identity
NCBI Description (AF022460) CYP83D1p [Glycine max]
Seq. No.
                  1214
                  647 2.R1040
Contig ID
5'-most EST
                  uC-gmronoir004e08b1
Method
                  BLASTN
                  g2739001
NCBI GI
BLAST score
                  145
                  2.0e-75
E value
Match length
                  249
% identity
                  90
NCBI Description Glycine max cytochrome P450 monooxygenase CYP83D1p
                  (CYP83D1) mRNA, partial cds
Seq. No.
                  1215
                  648 1.R1040
Contig ID
5'-most EST
                  jC-gmle01810086g05a1
                  BLASTX
Method
NCBI GI
                  q2739004
BLAST score
                  2805
E value
                  0.0e+00
Match length
                  532
% identity
                  100
NCBI Description (AF022461) CYP82Clp [Glycine max]
Seq. No.
                  1216
Contig ID
                  649 1.R1040
5'-most EST
                  txt700733671.h1
                  BLASTX
Method
                  q2739006
NCBI GI
BLAST score
                  2248
                  0.0e + 00
E value
Match length
                  498
                  89
% identity
NCBI Description (AF022462) CYP93Clp [Glycine max]
Seq. No.
                  1217
                  649 2.R1040
Contig ID
5'-most EST
                  g5230725_FL
Method
                  BLASTX
NCBI GI
                  g2739006
BLAST score
                  2152
E value
                  0.0e + 00
```



```
Seq. No.
                  649 3.R1040
Contiq ID
5'-most EST
                  jC-qmst02400030a06d1
                  BLASTX
Method
NCBI GI
                  g2739371
BLAST score
                  476
E value
                  1.0e-47
Match length
                  141
                  70
% identity
NCBI Description (AC002505) unknown protein [Arabidopsis thaliana]
Seq. No.
Contig ID
                  649 4.R1040
5'-most EST
                  LIB3138-019-Q1-N1-E7
Seq. No.
                  1220
Contig ID
                  649 7.R1040
                  jC-gmle01810071q04a1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q2739005
BLAST score
                  370
                  0.0e+00
E value
                  525
Match length
% identity
                  93
NCBI Description Glycine max cytochrome P450 monooxygenase CYP93C1p
                  (CYP93C1) mRNA, complete cds
                  1221
Seq. No.
Contig ID
                  649 8.R1040
                  sat701015083.hl
5'-most EST
Method
                  BLASTN
NCBI GI
                  g2739005
BLAST score
                  121
                  1.0e-61
E value
Match length
                  189
                  92
% identity
NCBI Description
                  Glycine max cytochrome P450 monooxygenase CYP93C1p
                   (CYP93C1) mRNA, complete cds
Seq. No.
                  1222
Contig ID
                   651 1.R1040
5'-most EST
                  ncj700984140.h1
Method
                  BLASTX
NCBI GI
                  q2739010
BLAST score
                  2560
E value
                  0.0e+00
Match length
                  513
                  96
% identity
NCBI Description (AF022464) CYP77A3p [Glycine max]
                  1223
Seq. No.
Contig ID
                   653 1.R1040
```

5'-most EST g2739043 FL

Method BLASTN NCBI GI g2739043 BLAST score 1422 E value 0.0e + 00



Match length 1426 100 % identity

Glycine max polyphosphoinositide binding protein Sshlp NCBI Description

(SSH1) mRNA, complete cds

Seq. No. 1224

653 3.R1040 Contig ID leu701149472.h1 5'-most EST

Method BLASTN NCBI GI q2739043 BLAST score 128 E value 8.0e-66 Match length 153 96 % identity

Glycine max polyphosphoinositide binding protein Ssh1p NCBI Description

(SSH1) mRNA, complete cds

1225 Seq. No.

Contig ID 653 5.R1040

LIB3028-044-Q1-B1-G1 5'-most EST

Method BLASTN g2739043 NCBI GI 139 BLAST score 2.0e-72 E value 209 Match length % identity 91

NCBI Description Glycine max polyphosphoinositide binding protein Sshlp

(SSH1) mRNA, complete cds

1226 Seq. No.

655 3.R1040 Contig ID

uC-gmropic059d06b1 5'-most EST

BLASTN Method g3169543 NCBI GI BLAST score 461 0.0e+00E value 2275 Match length 84 % identity

NCBI Description Medicago sativa sucrose synthase mRNA, complete cds

Seq. No.

1227

655 4.R1040 Contig ID

jC-gmro02910073c10a1 5'-most EST

BLASTX Method g3915046 NCBI GI 679 BLAST score 0.0e + 00E value 791 Match length % identity

SUCROSE SYNTHASE (SUCROSE-UDP GLUCOSYLTRANSFERASE) NCBI Description

>gi_3169544 (AF049487) sucrose synthase [Medicago sativa]

1228 Seq. No.

657 1.R1040 Contig ID g2654093_FL 5'-most EST Method BLASTX g2654094 NCBI GI



BLAST score 0.0e + 00E value Match length 456 93 % identity

(AF034210) aspartate aminotransferase glyoxysomal isozyme NCBI Description

AAT1 precursor [Glycine max]

1229 Seq. No.

658 1.R1040 Contig ID

LIB3049-014-Q1-E1-B12 5'-most EST

BLASTN Method q3377793 NCBI GI 929 BLAST score 0.0e+00E value 965 Match length 99 % identity

NCBI Description Glycine max proteasome IOTA subunit mRNA, complete cds

1230 Seq. No.

658 2.R1040 Contig ID g4291492 5'-most EST BLASTN Method g3377793 NCBI GI 330 BLAST score 0.0e + 00E value 390 Match length

96 % identity

NCBI Description Glycine max proteasome IOTA subunit mRNA, complete cds

Seq. No. 1231

659 1.R1040 Contig ID

LIB3030-005-Q1-B1-A12 5'-most EST

BLASTX Method NCBI GI q231695 2683 BLAST score E value 0.0e+00Match length 492 100 % identity

NCBI Description CATALASE >gi_66299_pir__CSSY catalase (EC 1.11.1.6) -

soybean >qi 18560 emb CAA78056 (Z12021) catalase [Glycine

 \max] >gi $26\overline{6}1015$ (AF0 $\overline{3}5252$) catalase [Glycine max]

>gi 2661 $\overline{0}$ 17 (AF035253) catalase [Glycine max]

1232 Seq. No.

659 2.R1040 Contig ID

LIB3051-080-Q1-K1-F12 5'-most EST

BLASTN Method g2661014 NCBI GI 121 BLAST score 1.0e-61 E value 167 Match length 96 % identity

NCBI Description Glycine max catalase (cat1) mRNA, complete cds

1233 Seq. No.

659 4.R1040 Contig ID

LIB3107-016-Q1-K1-G1 5'-most EST



```
BLASTN
Method
                  g2661018
NCBI GI
BLAST score
                  294
                  1.0e-164
E value
Match length
                  302
                  99
% identity
NCBI Description Glycine max catalase (cat3) mRNA, complete cds
                  1234
Seq. No.
                   659 6.R1040
Contig ID
                   zsg701119529.hl
5'-most EST
Method
                  BLASTN
                   g18559
NCBI GI
                   258
BLAST score
                   1.0e-143
E value
Match length
                   390
% identity
                   98
NCBI Description G.max gene for catalase
                   1235
Seq. No.
                   662_1.R1040
Contig ID
                   \verb"jC-gmle01810020g04a1"
5'-most EST
                   BLASTX
Method
                   g2661021
NCBI GI
BLAST score
                   2689
                   0.0e + 00
E value
Match length
                   492
                   100
% identity
NCBI Description (AF035255) catalase [Glycine max]
                   1236
Seq. No.
                   662 2.R1040
Contig ID
                   fC-gmst700889880r1
5'-most EST
                   BLASTN
Method
                   g2661022
NCBI GI
                   388
BLAST score
                   0.0e+00
E value
                   388
Match length
                   100
 % identity
 NCBI Description Glycine max catalase (cat5) mRNA, partial cds
                   1237
 Seq. No.
                   662 3.R1040
 Contig ID
                   LIB3051-025-Q1-K1-A8
 5'-most EST
                   BLASTN
 Method
                   g2661020
 NCBI GI
                    241
 BLAST score
                    1.0e-133
 E value
                    337
 Match length
                    99
 % identity
 NCBI Description Glycine max catalase (cat4) mRNA, complete cds
```

MCBI Description

Seq. No.

1230

Contig ID 662_5.R1040

5'-most EST jC-gmle01810082c02d1

Method BLASTN NCBI GI g2661020

305



```
BLAST score
                   337
                   0.0e+00
E value
                  341
Match length
                  100
% identity
                  Glycine max catalase (cat4) mRNA, complete cds
NCBI Description
                   1239
Seq. No.
                   662 8.R1040
Contig ID
5'-most EST
                   awf700838945.hl
                  BLASTN
Method
NCBI GI
                   q2661020
BLAST score
                   175
E value
                   6.0e-94
Match length
                   251
                   97
% identity
NCBI Description Glycine max catalase (cat4) mRNA, complete cds
                   1240
Seq. No.
                   664 1.R1040
Contig ID
5'-most EST
                   asn701139293.h1
                   BLASTX
Method
NCBI GI
                   q3334122
                   1767
BLAST score
                   0.0e + 00
E value
                   350
Match length
% identity
                  ARGINASE >gi_2661128 (AF035671) arginase [Glycine max]
NCBI Description
Seq. No.
                   1241
                   664 2.R1040
Contig ID
5'-most EST
                   kl1701203871.hl
                   BLASTX
Method
NCBI GI
                   g1168493
BLAST score
                   1464
E value
                   1.0e-163
Match length
                   330
% identity
                   87
                  ARGINASE >gi 602422 (U15019) arginase [Arabidopsis
NCBI Description
                   thaliana] >gi 4325373 gb AAD17369 (AF128396) Arabidopsis
                   thaliana arginase (SW:P46637) (Pfam: PF00491, Score=419.6,
                   E=3.7e-142 N=1) [Arabidopsis thaliana]
                   1242
Seq. No.
                   665 1.R1040
Contig ID
5'-most EST
                   zhf700960949.h1
                   BLASTX
Method
NCBI GI
                   q4104561
BLAST score
                   1689
                   0.0e+00
E value
```

NCBI Description

Match length % identity

1243

370

Seq. No. 666 1.R1040 Contig ID g5002233_FL 5'-most EST Method BLASTN

(AF036960) subtilisin-like protease [Glycine max]



```
NCBI GI
                  q3982595
BLAST score
                  884
E value
                  0.0e+00
Match length
                  1252
% identity
                  95
NCBI Description Glycine max cationic peroxidase 2 (Prx2) mRNA, complete cds
Seq. No.
Contig ID
                  666 2.R1040
5'-most EST
                  LIB3109-045-Q1-K1-A9
                  BLASTX
Method
NCBI GI
                  g3982596
BLAST score
                  1680
E value
                  0.0e+00
                  331
Match length
                  96
% identity
NCBI Description (AF039027) cationic peroxidase 2 [Glycine max]
Seq. No.
                  1245
                  666 3.R1040
Contig ID
                  LIB3107-063-Q1-K1-A3
5'-most EST
                  1246
Seq. No.
Contig ID
                  666 4.R1040
5'-most EST
                  asn701138013.hl
Method
                  BLASTN
                  g3982595
NCBI GI
BLAST score
                  67
                  3.0e-29
E value
                  250
Match length
                  87
% identity
NCBI Description Glycine max cationic peroxidase 2 (Prx2) mRNA, complete cds
                  1247
Seq. No.
                   667 1.R1040
Contig ID
                   jC-gmle01810025e03a1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q2773270
BLAST score
                   5249
E value
                  0.0e + 00
Match length
                  1023
                  100
% identity
                  (AF039851) ferredoxin-dependent glutamate synthase [Glycine
NCBI Description
                  max]
                  1248
Seq. No.
                  667 2.R1040
Contig ID
5'-most EST
                  bth700845893.h1
Method
                  BLASTX
                  q2494793
NCBI GI
BLAST score
                  689
```

NCBI GI g249479
BLAST score 689
E value 8.0e-73
Match length 159
% identity 82

NCBI Description [Segment 2 of 2] FERREDOXIN-DEPENDENT GLUTAMATE SYNTHASE (FD-GOGAT) >gi_2146724_pir__S67496 ferredoxin-dependent glutamate synthase (EC 1.4.7.1) - spinach (fragment)



>gi_414979 (U03006) ferredoxin-dependent glutamate synthase [Spinacia oleracea]

Seq. No. 1249

Contig ID 668 1.R1040

5'-most EST LIB3052-015-Q1-N1-A6

Method BLASTX
NCBI GI g120680
BLAST score 233
E value 2.0e-19
Match length 53
% identity 79

NCBI Description GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC

>gi 66010 pir DEBHG glyceraldehyde-3-phosphate

dehydrogenase (EC 1.2.1.12) - barley

>gi_18978_emb_CAA42901_ (X60343) glyceraldehyde 3-phosphate

dehydrogenase [Hordeum vulgare]

Seq. No. 1250

Contig ID 668 2.R1040

5'-most EST LIB3139-002-Q1-N1-B6

Method BLASTX
NCBI GI g462137
BLAST score 589
E value 1.0e-60
Match length 186
% identity 68

NCBI Description GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC

Seq. No. 1251

Contig ID 668_3.R1040

5'-most EST LIB3028-045-Q1-B1-D7

Method BLASTN
NCBI GI g169090
BLAST score 160
E value 2.0e-84
Match length 466
% identity 86

NCBI Description Pea glyceraldehyde-3-phosphate dehydrogenase (GapC1) mRNA,

complete cds

Seq. No. 1252

Contig ID 668_4.R1040 5'-most EST kl1701209936.h1

Method BLASTN
NCBI GI g169090
BLAST score 186
E value 1.0e-100
Match length 477
% identity 85

NCBI Description Pea glyceraldehyde-3-phosphate dehydrogenase (GapC1) mRNA,

complete cds

Seq. No. 1253

Contig ID 668_5.R1040

5'-most EST LIB3087-006-Q1-K1-G1

Method BLASTN



NCBI GI g166705 BLAST score 224 E value 1.0e-122 Match length 951 % identity 83

NCBI Description Arabidopsis thaliana glyceraldehyde-3-phosphate dehydrogenase C subunit (GapC) gene, complete cds

Seq. No. 1254

Contig ID 668 6.R1040

5'-most EST jC-gmf102220103a05a1

Method BLASTN
NCBI GI g169090
BLAST score 262
E value 1.0e-145
Match length 698
% identity 85

NCBI Description Pea glyceraldehyde-3-phosphate dehydrogenase (GapC1) mRNA,

complete cds

Seq. No. 1255

Contig ID 670 1.R1040

5'-most EST LIB3139-017-P1-N1-H1

Method BLASTX
NCBI GI g4056457
BLAST score 343
E value 1.0e-53
Match length 135
% identity 81

NCBI Description (AC005990) ESTs gb 234051 and gb F13722 come from this

gene. [Arabidopsis thaliana]

Seq. No. 1256

Contig ID 670 2.R1040

5'-most EST LIB3073-025-Q1-K1-E5

Method BLASTN
NCBI GI g2905777
BLAST score 182
E value 8.0e-98
Match length 266
% identity 92

NCBI Description Glycine max ribosomal protein L41 mRNA, partial cds

Seq. No. 1257

Contig ID 670_3.R1040 5'-most EST sat701010357.h1

Method BLASTN
NCBI GI g2905777
BLAST score 239
E value 1.0e-131
Match length 340
% identity 98

NCBI Description Glycine max ribosomal protein L41 mRNA, partial cds

Seq. No. 1258

Contig ID 670_4.R1040

5'-most EST uC-gmropic018h03b1



```
BLASTN
Method
                  g2905777
NCBI GI
BLAST score
                  167
                  7.0e-89
E value
Match length
                  243
                  92
% identity
NCBI Description Glycine max ribosomal protein L41 mRNA, partial cds
                  1259
Seq. No.
                   670 5.R1040
Contig ID
                  LIB3049-044-Q1-E1-C3
5'-most EST
                   BLASTN
Method
                   g2905777
NCBI GI
BLAST score
                   160
                   1.0e-84
E value
Match length
                   244
% identity
                   92
                  Glycine max ribosomal protein L41 mRNA, partial cds
NCBI Description
                   1260
Seq. No.
                   670_7.R1040
Contig ID
                   q5605749
5'-most EST
                   BLASTN
Method
                   g2905777
NCBI GI
BLAST score
                   167
                   1.0e-88
E value
                   243
Match length
                   92
% identity
NCBI Description Glycine max ribosomal protein L41 mRNA, partial cds
                   1261
Seq. No.
                   670 10.R1040
Contig ID
                   epx701109978.h1
5'-most EST
                   1262
Seq. No.
                   671 1.R1040
Contig ID
                   LIB3170-056-Q1-K1-B7
5'-most EST
                   1263
 Seq. No.
                   671 2.R1040
Contig ID
                   q4397284
 5'-most EST
                   BLASTN
Method
                   ģ2905779
 NCBI GI
                   257
 BLAST score
                   1.0e-142
 E value
                   343
 Match length
 % identity
                   Glycine max hydroxyproline-rich glycoprotein (hrgp) mRNA,
 NCBI Description
                   partial cds
                    1264
 Seq. No.
                    671 3.R1040
 Contig ID
                    a4396481
 5'-most EST
```

Contig ID 671_3.R1040
5'-most EST g4396481
Method BLASTN
NCBI GI g2905779
BLAST score 163
E value 2.0e-86



Match length 206 % identity 96

NCBI Description Glycine max hydroxyproline-rich glycoprotein (hrgp) mRNA,

partial cds

Seq. No. 1265

Contig ID 673_1.R1040

5'-most EST LIB3170-074-Q1-K1-G8

Seq. No. 1266

Contig ID 673 2.R1040

5'-most EST LIB3073-020-Q1-K1-F7

Seq. No. 1267

673 3.R1040 Contig ID 5'-most EST g3005575 FL Method BLASTX NCBI GI q3005576 BLAST score 2576 E value 0.0e + 00Match length 516 97 % identity

NCBI Description (AF047718) putative high affinity nitrate transporter;

GmNRT2 [Glycine max]

Seq. No. 1268

Contig ID 673 4.R1040

5'-most EST jC-gmro02800043e01a1

Method BLASTN
NCBI GI g3005575
BLAST score 310
E value 1.0e-174
Match length 370
% identity 96

NCBI Description Glycine max putative high affinity nitrate transporter

(NRT2) mRNA, complete cds

Seq. No. 1269

Contig ID 678_1.R1040 5'-most EST vwf700676019.h1

Method BLASTN
NCBI GI g2827079
BLAST score 351
E value 0.0e+00
Match length 803
% identity 86

NCBI Description Medicago sativa mitochondrial malate dehydrogenase

precursor (mmdh) mRNA, nuclear gene encoding mitochondrial

protein, complete cds

Seq. No. 1270

Contig ID 678_3.R1040 5'-most EST kl1701208643.h1

Method BLASTN
NCBI GI g2827079
BLAST score 63
E value 5.0e-27



Match length 115 % identity 89

NCBI Description Medicago sativa mitochondrial malate dehydrogenase

precursor (mmdh) mRNA, nuclear gene encoding mitochondrial

protein, complete cds

Seq. No. 1271

Contig ID 678_4.R1040 5'-most EST zsg701125938.h1

Method BLASTN
NCBI GI g2827079
BLAST score 44
E value 1.0e-15
Match length 96
% identity 86

NCBI Description Medicago sativa mitochondrial malate dehydrogenase

precursor (mmdh) mRNA, nuclear gene encoding mitochondrial

protein, complete cds

Seq. No. 1272

Contig ID 679_1.R1040 5'-most EST pxt700942053.h1

Method BLASTN
NCBI GI g3193221
BLAST score 318
E value 1.0e-179
Match length 516
% identity 93

NCBI Description Glycine max malate dehydrogenase (Mdh-2) gene, nuclear gene

encoding mitochondrial protein, partial cds

Seq. No. 1273

Contig ID 680_1.R1040

5'-most EST uC-gmflminsoy067d03b1

Method BLASTX
NCBI GI g3399777
BLAST score 1700
E value 0.0e+00
Match length 347
% identity 97

NCBI Description (AF069738) symbiotic ammonium transporter; nodulin [Glycine

max]

Seq. No. 1274

Contig ID 682_1.R1040 5'-most EST zhf700956708.h1

Method BLASTN
NCBI GI g3552012
BLAST score 1260
E value 0.0e+00
Match length 1304
% identity 99

NCBI Description Glycine max proline-rich protein precursor, mRNA, complete

cds

Seq. No. 1275

Contig ID 682 3.R1040



```
5'-most EST
                   jC-gmst02400067c05a1
                  BLASTN
Method
NCBI GI
                  g3552012
BLAST score
                  145
E value
                  1.0e-75
                  504
Match length
                  88
% identity
                  Glycine max proline-rich protein precursor, mRNA, complete
NCBI Description
                  cds
                  1276
Seq. No.
Contig ID
                  683 1.R1040
5'-most EST
                  g957422
                  BLASTX
Method
NCBI GI
                  g1076511
BLAST score
                  2415
E value
                  0.0e+00
Match length
                  522
                  90
% identity
                  H+-transporting ATPase (EC 3.6.1.35) - kidney bean
NCBI Description
                  >qi 758250 emb CAA59799 (X85804) H(+)-transporting ATPase
                  [Phaseolus vulgaris]
Seq. No.
                  1277
                  683 2.R1040
Contig ID
                  jC-gmst02400039g11a1
5'-most EST
                  BLASTN
Method
                  g3644031
NCBI GI
BLAST score
                  175
E value
                  2.0e-93
                  183
Match length
                  99
% identity
NCBI Description Glycine max plasma membrane proton-ATPase mRNA, partial cds
                   1278
Seq. No.
Contig ID
                   683 3.R1040
                  LIB3051-064-Q1-K1-G2
5'-most EST
Method
                  BLASTX
NCBI GI
                  q1076511
BLAST score
                  1765
E value
                  0.0e+00
                  371
Match length
                  94
% identity
                  H+-transporting ATPase (EC 3.6.1.35) - kidney bean
NCBI Description
                  >gi 758250 emb_CAA59799 (X85804) H(+)-transporting ATPase
                   [Phaseolus vulgaris]
Seq. No.
                   1279
Contig ID
                   683 5.R1040
5'-most EST
                   jC-gmf102220060a11d1
Method
                  BLASTN
NCBI GI
                  g3644031
BLAST score
                  180
E value
                  1.0e-96
                  180
Match length
```

100

% identity



81

% identity

```
Seq. No.
                  1280
                  683 6.R1040
Contig ID
                  uC-gmflminsoy037a02b1
5'-most EST
                  BLASTN
Method
NCBI GI
                  g758249
BLAST score
                  200
                  1.0e-108
E value
                  396
Match length
                  88
% identity
NCBI Description P.vulgaris mRNA for plasma membrane H+ ATPase
                  1281
Seq. No.
Contig ID
                  683 8.R1040
5'-most EST
                  q5508913
Method
                  BLASTN
NCBI GI
                  g758249
BLAST score
                  116
                  1.0e-58
E value
                  254
Match length
% identity
                  94
NCBI Description P.vulgaris mRNA for plasma membrane H+ ATPase
                  1282
Seq. No.
Contig ID
                  684 1.R1040
5'-most EST
                  uC-gmflminsoy014b09b1
                  BLASTX
Method
NCBI GI
                  q2275219
BLAST score
                  1852
                  0.0e+00
E value
                  418
Match length
                  83
% identity
NCBI Description (AC002337) unknown protein [Arabidopsis thaliana]
Seq. No.
                  1283
                  684 2.R1040
Contig ID
5'-most EST
                  LIB3051-006-Q1-E1-F11
                  BLASTX
Method
                  g2275219
NCBI GI
                  371
BLAST score
E value
                  1.0e-35
Match length
                  138
% identity
                  61
NCBI Description (AC002337) unknown protein [Arabidopsis thaliana]
Seq. No.
                  1284
Contig ID
                  685 1.R1040
5'-most EST
                  rlr700895815.h1
Method
                  BLASTX
                  g123506
NCBI GI
BLAST score
                  336
E value
                  4.0e-31
Match length
                  80
```

hydrophobic seed protein - soybean >gi_349888 pdb_1HYP_

NCBI Description HYDROPHOBIC SEED PROTEIN (HPS) >gi 99917 pir A29385

Hydrophobic Protein From Soybean



Seq. No. 1285

Contig ID 685_2.R1040 5'-most EST rlr700895643.h1

Method BLASTX
NCBI GI g123506
BLAST score 249
E value 3.0e-21
Match length 66
% identity 74

NCBI Description HYDROPHOBIC SEED PROTEIN (HPS) >gi 99917_pir A29385

hydrophobic seed protein - soybean >gi_349888 pdb_1HYP_

Hydrophobic Protein From Soybean

Seq. No. 1286

Contig ID 685_3.R1040 5'-most EST rlr700897394.h1

Seq. No. 1287

Contig ID 686_1.R1040 5'-most EST uxk700672239.h1

Method BLASTN
NCBI GI g3832527
BLAST score 1500
E value 0.0e+00
Match length 1521
% identity 100

NCBI Description Glycine max unknown mRNA

Seq. No. 1288

Contig ID 686_2.R1040

5'-most EST uC-gmflminsoy042g06b1

Method BLASTN
NCBI GI g3832527
BLAST score 73
E value 8.0e-33
Match length 73
% identity 100

NCBI Description Glycine max unknown mRNA

Seq. No. 1289

Contig ID 689 3.R1040 5'-most EST eep700863946.h1

Seq. No.

1290

Contig ID 689_4.R1040

5'-most EST uC-gmrominsoy081g04b1

Method BLASTX
NCBI GI g3831470
BLAST score 184
E value 2.0e-13
Match length 92
% identity 49

NCBI Description (AC005700) unknown protein, 3' partial [Arabidopsis

thaliana]

Seq. No.

1291



```
Contig ID
                  689 5.R1040
                  LIB3050-017-Q1-E1-C9
5'-most EST
Method
                  BLASTX
                  g3851670
NCBI GI
BLAST score
                  285
E value
                  3.0e-25
Match length
                  111
% identity
                  56
NCBI Description (AF099906) J8 [Arabidopsis thaliana]
                  1292
Seq. No.
Contig ID
                  689 6.R1040
5'-most EST
                  LIB3051-011-Q1-E1-A7
Method
                  BLASTX
                  g1388078
NCBI GI
BLAST score
                  244
E value
                  9.0e-22
Match length
                  96
% identity
NCBI Description (U35826) thioredoxin h [Arabidopsis thaliana]
                  1293
Seq. No.
                  689 7.R1040
Contig ID
5'-most EST
                  LIB3049-029-Q1-E1-G7
Method
                  BLASTX
NCBI GI
                  g4455358
                  487
BLAST score
E value
                  8.0e-49
Match length
                  184
% identity
                  61
NCBI Description (AL035524) putative protein [Arabidopsis thaliana]
Seq. No.
                  1294
                  689 8.R1040
Contig ID
5'-most EST
                  LIB3055-008-Q1-N1-B9
Method
                  BLASTX
NCBI GI
                  g3269293
BLAST score
                  272
                  2.0e-49
E value
                  213
Match length
% identity
NCBI Description (AL030978) putative protein [Arabidopsis thaliana]
                  1295
Seq. No.
                  689 9.R1040
Contig ID
5'-most EST
                  jC-gmst02400052b03a1
Method
                  BLASTX
NCBI GI
                  g3851670
BLAST score
                  281
E value
                  1.0e-24
```

Match length 112 % identity 51

NCBI Description (AF099906) J8 [Arabidopsis thaliana]

1296 Seq. No.

689 10.R1040 Contig ID

5'-most EST jC-gmst02400004g01a1



Method BLASTX
NCBI GI g3269293
BLAST score 472
E value 6.0e-47
Match length 179
% identity 56

NCBI Description (AL030978) putative protein [Arabidopsis thaliana]

Seq. No.

1297

Contig ID 5'-most EST

689_11.R1040 rlr700900707.h1

Method BLASTN
NCBI GI g2687435
BLAST score 449
E value 0.0e+00
Match length 645
% identity 93

NCBI Description Hamamelis virginiana large subunit 26S ribosomal RNA gene,

partial sequence

Seq. No.

1298

Contig ID

689 12.R1040

5'-most EST

LIB3050-008-Q1-E1-E9

Method BLASTX
NCBI GI g2129752
BLAST score 437
E value 5.0e-43
Match length 104
% identity 75

NCBI Description thioredoxin - Arabidopsis thaliana >gi 992964 emb CAA84612_

(Z35475) thioredoxin [Arabidopsis thaliana]

Seq. No.

1299

Contig ID

689 13.R1040

5'-most EST

LIB3170-046-Q1-J1-B4

Method BLASTN
NCBI GI g2852442
BLAST score 224
E value 1.0e-122
Match length 510
% identity 86

NCBI Description Salix bakko mRNA for ADP-ribosylation factor, complete cds

Seq. No.

1300

Contig ID

689 14.R1040

5'-most EST

jC-gmro02910071a04a1

Method BLASTX
NCBI GI g1708236
BLAST score 1265
E value 0.0e+00
Match length 462
% identity 66

NCBI Description HYDROXYMETHYLGLUTARYL-COA SYNTHASE (HMG-COA SYNTHASE)

(3-HYDROXY-3-METHYLGLUTARYL COENZYME A SYNTHASE)

>gi_2129617_pir__JC4567 hydroxymethylglutaryl-CoA synthase

(EC 4.1.3.5) - Arabidopsis thaliana >gi 1143390 emb CAA58763 (X83882)



hydroxymethylglutaryl-CoA synthase [Arabidopsis thaliana] >gi_1586548_prf__2204245A hydroxy methylglutaryl CoA synthase [Arabidopsis thaliana]

Seq. No. 1301 Contig ID 689 15.R1040

5'-most EST uC-gmropic080b12b1

Method BLASTN
NCBI GI g2852442
BLAST score 221
E value 1.0e-120
Match length 485
% identity 86

NCBI Description Salix bakko mRNA for ADP-ribosylation factor, complete cds

Seq. No. 1302

Contig ID 689 16.R1040 5'-most EST leu701156547.h1

Method BLASTN
NCBI GI g2852442
BLAST score 224
E value 1.0e-122
Match length 480
% identity 87

NCBI Description Salix bakko mRNA for ADP-ribosylation factor, complete cds

Seq. No. 1303

Contig ID 689_17.R1040 5'-most EST jex700909370.h1

Method BLASTX
NCBI GI g1708236
BLAST score 1544
E value 1.0e-179
Match length 464
% identity 67

% identity 6

NCBI Description HYDROXYMETHYLGLUTARYL-COA SYNTHASE (HMG-COA SYNTHASE)

(3-HYDROXY-3-METHYLGLUTARYL COENZYME A SYNTHASE)

>gi_2129617_pir__JC4567 hydroxymethylglutaryl-CoA synthase

(EC 4.1.3.5) - Arabidopsis thaliana >gi_1143390_emb_CAA58763_ (X83882)

hydroxymethylglutaryl-CoA synthase [Arabidopsis thaliana]

>gi_1586548_prf__2204245A hydroxy methylglutaryl CoA

synthase [Arabidopsis thaliana]

Seq. No. 1304

Contig ID 689 18.R1040

5'-most EST LIB3050-023-Q1-K1-C8

Method BLASTN
NCBI GI g4324966
BLAST score 349
E value 0.0e+00
Match length 610 ***
% identity 96

NCBI Description Glycine max ADP-ribosylation factor mRNA, partial cds

Seq. No. 1305

Contig ID 690 1.R1040



crh700852726.h1 5'-most EST

BLASTN Method g472849 NCBI GI 180 BLAST score 2.0e-96 E value 460 Match length 85 % identity

Glycine max Essex desiccation protectant protein Lea14 NCBI Description

homolog mRNA, complete cds

1306 Seq. No.

691 1.R1040 Contig ID

LIB3051-047-Q1-K1-E6 5'-most EST

1307 Seq. No.

691 2.R1040 Contig ID 5'-most EST al1700863202.h1

1308 Seq. No.

692 1.R1040 Contig ID

LIB3027-008-Q1-B1-H3 5'-most EST

1309 Seq. No.

694 1.R1040 Contig ID

LIB3072-017-Q1-E1-E5 5'-most EST

BLASTX Method g2909422 NCBI GI 395 BLAST score 6.0e-38 E value Match length 135 % identity

(AJ224519) LEA PROTEIN [Cicer arietinum] NCBI Description

1310 Seq. No.

694 2.R1040 Contig ID

LIB3106-072-P1-K1-B2 5'-most EST

BLASTX Method g2245062 NCBI GI 185 BLAST score 2.0e-13 E value Match length 44 % identity 77

NCBI Description (Z97342) hypothetical protein [Arabidopsis thaliana]

1311 Seg. No.

694 4.R1040 Contig ID g5688020 5'-most EST

1312 Seq. No.

695 1.R1040 Contig ID crh700853905.hl 5'-most EST

Method BLASTX q169101 NCBI GI 407 BLAST score 2.0e-39 E value 127 Match length 62 % identity



```
(M33900) 17.9 kDa heat shock protein (hsp17.9) [Pisum
NCBI Description
                  sativum]
Seq. No.
                  1313
                  696 1.R1040
Contig ID
                  g4559291_FL
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4559292
BLAST score
                  2962
E value
                  0.0e + 00
                  557
Match length
                  100
% identity
NCBI Description (AF124148) trehalase 1 GMTRE1 [Glycine max]
Seq. No.
                  1314
                  698 1.R1040
Contig ID
5'-most EST
                  LIB3109-042-Q1-K1-D5
Method
                  BLASTX
NCBI GI
                  q4567091
BLAST score
                  2104
E value
                  0.0e + 00
Match length
                  507
% identity
                  96
                  (AF128443) SNF-1-like serine/threonine protein kinase
NCBI Description
                  [Glycine max]
Seq. No.
                  1315
Contig ID
                  698 2.R1040
5'-most EST
                  LIB3051-030-Q1-K1-G7
Method
                  BLASTN
NCBI GI
                  g4567090
BLAST score
                  105
                  1.0e-51
E value
Match length
                  205
% identity
                  91
NCBI Description
                 Glycine max SNF-1-like serine/threonine protein kinase
                  mRNA, complete cds
Seq. No.
                  1316
                  698 3.R1040
Contig ID
5'-most EST
                  jex700907708.hl
Seq. No.
                  1317
                  698 4.R1040
Contig ID
5'-most EST
                  seb700651919.hl
Method
                  BLASTN
NCBI GI
                  g4567090
BLAST score
                  137
                  6.0e-71
E value
Match length
                  277
                  87
% identity
NCBI Description
                  Glycine max SNF-1-like serine/threonine protein kinase
```

mRNA, complete cds

1318

698 5.R1040

zsg701128474.h1

Seq. No. Contig ID

5'-most EST



Method BLASTN NCBI GI q4567090 BLAST score 137 E value 7.0e-71 Match length 407 86 % identity

NCBI Description Glycine max SNF-1-like serine/threonine protein kinase

mRNA, complete cds

Seq. No. 1319 699 1.R1040 Contig ID xpa700793223.hl 5'-most EST

BLASTX Method q130846 NCBI GI BLAST score 494 1.0e-49 E value 161 Match length 57 % identity

BASIC FORM OF PATHOGENESIS-RELATED PROTEIN 1 PRECURSOR (PRP NCBI Description

> 1) >gi_82185_pir__S04728 pathogenesis-related protein homolog precursor - common tobacco >gi 19972 emb CAA32228 (X14065) PRP 1 precursor (AA -23 to 154) [Nicotiana

tabacum] >qi 228620 prf 1807333A pathogenesis-related

protein 1 [Nicotiana tabacum]

Seq. No. 1320

Contig ID 702 1.R1040 5'-most EST g5002235 FL Method BLASTX g4538927 NCBI GI BLAST score 1138 1.0e-125 E value Match length 302

% identity 68

(AL049483) putative peroxidase [Arabidopsis thaliana] NCBI Description

Seq. No. 1321

703 1.R1040 Contig ID zzp700834588.h1 5'-most EST

Method BLASTX NCBI GI q478409 1020 BLAST score E value 1.0e-111 320 Match length % identity 62

peroxidase (EC 1.11.1.7), cationic - adzuki bean NCBI Description

>gi 218328 dbj BAA01950 (D11337) peroxidase [Vigna

angularis]

Seq. No. 1322

Contig ID 704 1.R1040 5'-most EST zhf700962226.h1

Method BLASTX NCBI GI q4512617 BLAST score 255 E value 2.0e-21 Match length 157



% identity 44 NCBI Description (AC004793) F28K20.16 [Arabidopsis thaliana] Seq. No. 1323

Contig ID 705_1.R1040 5'-most EST sat701002940.h1 Method BLASTN NCBI GI g1938226 BLAST score 290

BLAST score 290 E value 1.0e-162 Match length 745 % identity 85

NCBI Description M.sativa mRNA for carbonic anhydrase

Seq. No. 1324

Contig ID 706 1.R1040

5'-most EST LIB3093-024-Q1-K1-B2

Method BLASTX
NCBI GI g3413473
BLAST score 1781
E value 0.0e+00
Match length 338
% identity 100

NCBI Description (AJ006308) tyrosine phosphatase 1 [Glycine max]

Seq. No. 1325

Contig ID 707_1.R1040

5'-most EST LIB3170-084-Q1-K1-A6

Method BLASTN
NCBI GI g533691
BLAST score 483
E value 0.0e+00
Match length 502
% identity 95

NCBI Description Glycine max Essex protease inhibitor mRNA, complete cds

Seq. No. 1326

Contig ID 708 1.R1040

5'-most EST LIB3170-086-Q1-J1-B5

Method BLASTX
NCBI GI g905361
BLAST score 3108
E value 0.0e+00
Match length 640
% identity 93

NCBI Description (U22103) gag-protease polyprotein [Glycine max]

Seq. No. 1327

Contig ID 710_1.R1040

5'-most EST LIB3170-007-Q1-K1-D8

Method BLASTX
NCBI GI g3915009
BLAST score 3105
E value 0.0e+00
Match length 691
% identity 87

NCBI Description ARGININE DECARBOXYLASE (ARGDC) (ADC) >qi 1163181 (U35367)

Match length

% identity

152

95



arginine decarboxylase [Glycine max]

```
Seq. No.
                  1328
                  710 2.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy165f06b1
Method
                  BLASTN
NCBI GI
                  g1163180
BLAST score
                  278
                  1.0e-155
E value
Match length
                  316
% identity
                  97
NCBI Description Glycine max arginine decarboxylase mRNA, complete cds
Seq. No.
Contig ID
                  715 1.R1040
5'-most EST
                  kl1701212945.h1
Method
                  BLASTX
NCBI GI
                  g2832629
BLAST score
                  342
E value
                  1.0e-31
Match length
                  221
% identity
                  35
                  (AL021711) 4-coumarate-CoA ligase - like [Arabidopsis
NCBI Description
                  thaliana]
                  1330
Seq. No.
Contig ID
                  715 2.R1040
5'-most EST
                  g1143318 FL
Method
                  BLASTN
NCBI GI
                  q1143318
BLAST score
                  1067
                  0.0e+00
E value
Match length
                  1129
% identity
                   99
NCBI Description Glycine max biotin carboxyl carrier protein precursor
                   (accB-1) mRNA, complete cds
Seq. No.
                  1331
                  716 1.R1040
Contig ID
5'-most EST
                  g1335861 FL
Method
                  BLASTX
NCBI GI
                  g1335862
BLAST score
                  8090
                  0.0e+00
E value
                  1656
Match length
                  96
% identity
NCBI Description (U42608) clathrin heavy chain [Glycine max]
Seq. No.
                  1332
                  717 1.R1040
Contig ID
5'-most EST
                  leu701156669.h1
Method
                  BLASTN
NCBI GI
                  g758693
BLAST score
                  88
E value
                  2.0e-41
```



NCBI Description Catharanthus roseus S-adenosyl-L-methionine decarboxylase proenzyme mRNA, complete cds

Seq. No. 1333 Contig ID 717 2.R1040 5'-most EST g1438880 FL Method BLASTX NCBI GI g1438881 BLAST score 1912 E value 0.0e + 00Match length 362 100 % identity

NCBI Description (U43839) GmCK2p [Glycine max]

Seq. No.

Contig ID 717 3.R1040

5'-most EST LIB3028-005-Q1-B1-B4

1334

Method BLASTX
NCBI GI g3913420
BLAST score 1508
E value 1.0e-168
Match length 346
% identity 83

NCBI Description S-ADENOSYLMETHIONINE DECARBOXYLASE PROENZYME (ADOMETDC)

(SAMDC) >gi 1421752 (U60592) S-adenosylmethionine

decarboxylase [Pisum sativum]

Seq. No. 1335

Contig ID 717_4.R1040 5'-most EST rca700998660.h1

Method BLASTX
NCBI GI g3913420
BLAST score 1495
E value 1.0e-166
Match length 347
% identity 81

NCBI Description S-ADENOSYLMETHIONINE DECARBOXYLASE PROENZYME (ADOMETDC)

(SAMDC) >gi 1421752 (U60592) S-adenosylmethionine

decarboxylase [Pisum sativum]

Seq. No. 1336

Contig ID 717 5.R1040 5'-most EST g1438878 FL Method BLASTX q1438879 NCBI GI BLAST score 1900 E value 0.0e + 00Match length 359 % identity 100

NCBI Description (U43838) GmCK1p [Glycine max]

Seq. No. 1337

Contig ID 717_6.R1040

5'-most EST LIB3051-043-Q1-K1-C4

Method BLASTX NCBI GI g3860274 BLAST score 467



E value 2.0e-46 Match length 158 % identity 57

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

>gi_4314397_gb_AAD15607_ (AC006232) putative zinc finger

protein [Arabidopsis thaliana]

Seq. No. 1338

Contig ID 717_7.R1040 5'-most EST rca701002512.h1

Method BLASTN
NCBI GI g1438878
BLAST score 303
E value 1.0e-170
Match length 360
% identity 100

NCBI Description Glycine max choline kinase GmCK1p mRNA, complete cds

Seq. No. 1339

717 8.R1040 Contig ID 5'-most EST q5606823 Method BLASTX NCBI GI q3860274 BLAST score 216 E value 3.0e-17 Match length 54 % identity 74

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

>gi 4314397 gb AAD15607 (AC006232) putative zinc finger

protein [Arabidopsis thaliana]

Seq. No. 1340

Contig ID 717_9.R1040 5'-most EST epx701105723.h1

Method BLASTN
NCBI GI g1438878
BLAST score 430
E value 0.0e+00
Match length 448
% identity 99

NCBI Description Glycine max choline kinase GmCK1p mRNA, complete cds

Seq. No. 1341

717 10.R1040 Contig ID 5'-most EST g5057990 BLASTX Method NCBI GI g3860274 BLAST score 410 E value 7.0e-40 Match length 147 % identity 54

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

>gi_4314397_gb_AAD15607_ (AC006232) putative zinc finger

protein [Arabidopsis thaliana]

Seq. No. 1342

Contig ID 717 11.R1040



```
uC-gmronoir054h11b1
5'-most EST
                  BLASTX
Method
                  g3860274
NCBI GI
                  292
BLAST score
E value
                  4.0e-26
Match length
                  118
% identity
                  (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
                  >qi 4314397 gb AAD15607 (AC006232) putative zinc finger
                  protein [Arabidopsis thaliana]
Seq. No.
                  717 12.R1040
Contig ID
                  rca700997186.hl
5'-most EST
Method
                  BLASTN
NCBI GI
                  q1438878
BLAST score
                  345
E value
                   0.0e + 00
                  345
Match length
                  100
% identity
                  Glycine max choline kinase GmCK1p mRNA, complete cds
NCBI Description
                   1344
Seq. No.
                   717 14.R1040
Contig ID
                   jC-gmro02910026f07a1
5'-most EST
                   BLASTN
Method
NCBI GI
                   q1438880
                   256
BLAST score
                   1.0e-142
E value
Match length
                   372
                   95
% identity
                  Glycine max choline kinase GmCK2p mRNA, complete cds
NCBI Description
                   1345
Seq. No.
                   717 21.R1040
Contig ID
                   jC-gmst02400055f08d1
5'-most EST
Method
                   BLASTN
                   g1134881
NCBI GI
BLAST score
                   33
                   6.0e-09
E value
Match length
                   92
                   92
% identity
                  P.sativum mRNA for cysteine protease
NCBI Description
                   1346
Seq. No.
                   717 23.R1040
Contig ID
                   LIB3106-102-Q1-K1-E2
5'-most EST
                   BLASTN
Method
                   g758693
NCBI GI
```

Method BLASTN
NCBI GI 9758693
BLAST score 75
E value 6.0e-34
Match length 147
% identity 93

1347

NCBI Description Catharanthus roseus S-adenosyl-L-methionine decarboxylase

proenzyme mRNA, complete cds

Seq. No.

_

326



```
Contig ID
                  717 26.R1040
                  jC-gmro02910026d07d1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q1438880
BLAST score
                  272
E value
                  1.0e-151
Match length
                  363
% identity
                  96
NCBI Description Glycine max choline kinase GmCK2p mRNA, complete cds
Seq. No.
                  1348
Contig ID
                  719 1.R1040
5'-most EST
                  g1438882 FL
Method
                  BLASTX
NCBI GI
                  g1438883
BLAST score
                  2586
E value
                  0.0e + 00
Match length
                  497
% identity
                  100
NCBI Description
                  (U43840) GmCK3p [Glycine max]
Seq. No.
                  1349
Contig ID
                  722 1.R1040
5'-most EST
                  LIB3109-042-Q1-K1-G7
Method
                  BLASTX
NCBI GI
                  q4204759
BLAST score
                  1722
E value
                  0.0e+00
Match length
                  354
% identity
                  95
                  (U51191) peroxidase precursor [Glycine max]
NCBI Description
                  1350
Seq. No.
                  724 1.R1040
Contig ID
5'-most EST
                  LIB3107-052-Q1-K1-H12
Method
                  BLASTN
NCBI GI
                  g4204762
BLAST score
                  1030
E value
                  0.0e+00
Match length
                  1135
% identity
                  98
NCBI Description Glycine max peroxidase (sEPb1) mRNA, partial cds
                  1351
                  724 2.R1040
Contig ID
5'-most EST
                  g4204764_FL
Method
                  BLASTX
```

Seq. No.

NCBI GI g4204765 BLAST score 1671 E value 0.0e + 00Match length 325 % identity 100

NCBI Description (U51194) peroxidase [Glycine max]

Seq. No. 1352

Contig ID 726 1.R1040 5'-most EST ary700764391.h1

327

1



Method BLASTX
NCBI GI g1518540
BLAST score 2422
E value 0.0e+00
Match length 480
% identity 97

NCBI Description (U53418) UDP-glucose dehydrogenase [Glycine max]

Seq. No. 1353

Contig ID 726_2.R1040

5'-most EST jC-gmst02400027a06d1

Method BLASTN
NCBI GI g1518539
BLAST score 109
E value 3.0e-54
Match length 197
% identity 89

NCBI Description Glycine max UDP-glucose dehydrogenase mRNA, complete cds

Seq. No. 1354

Contig ID 726_6.R1040 5'-most EST jsh701069261.h1

Method BLASTX
NCBI GI g1518540
BLAST score 329
E value 5.0e-47
Match length 142
% identity 75

NCBI Description (U53418) UDP-glucose dehydrogenase [Glycine max]

Seq. No. 1355

Contig ID 727_2.R1040 5'-most EST taw700655529.h1

Method BLASTN
NCBI GI g1778371
BLAST score 158
E value 1.0e-83
Match length 220
% identity 99

NCBI Description Glycine max asparagine synthetase 1 (AS1) mRNA, complete

cds

Seq. No. 1356

Contig ID 727_3.R1040

5'-most EST fC-gmf1700906171a1

Method BLASTN
NCBI GI g1778371
BLAST score 366
E value 0.0e+00
Match length 927
% identity 97

NCBI Description Glycine max asparagine synthetase 1 (AS1) mRNA, complete

cds

Seq. No. 1357

Contig ID 727_4.R1040 5'-most EST crh700855008.h1



```
Method
                   BLASTX
NCBI GI
                   q3789911
BLAST score
                   143
E value
                   6.0e-27
Match length
                   192
% identity
                   43
NCBI Description
                   (AF081802) developmental protein DG1118 [Dictyostelium
                   discoideum]
                   1358
Seq. No.
Contig ID
                   727 5.R1040
5'-most EST
                   uC-gmropic067d04b1
Method
                   BLASTX
                   g1778372
NCBI GI
BLAST score
                   1198
E value
                   0.0e + 00
Match length
                   573
% identity
                  55
NCBI Description
                  (U77679) asparagine synthetase 1 [Glycine max]
                  1359
Seq. No.
Contig ID
                  727 6.R1040
5'-most EST
                  LIB3139-049-P1-N1-G6
Method
                  BLASTX
NCBI GI
                  g1778370
                  804
BLAST score
E value
                  0.0e + 00
Match length
                  579
% identity
                   42
NCBI Description
                  (U77678) asparagine synthetase 2 [Glycine max]
                  1360
Seq. No.
Contig ID
                  728 1.R1040
5'-most EST
                  leu701144891.h1
Method
                  BLASTN
NCBI GI
                  g1381675
BLAST score
                  739
E value
                  0.0e + 00
Match length
                  747
% identity
                  100
NCBI Description
                  Glycine max small GTP-binding protein (sra1) mRNA, partial
Seq. No.
                  1361
Contig ID
                  728 2.R1040
5'-most EST
                  jC-gmf102220078d07a1
Method
                  BLASTX
NCBI GI
                  q4262154
```

Method BLASTX
NCBI GI g426215
BLAST score 257
E value 1.0e-21
Match length 179
% identity 41

NCBI Description (AC005275) putative protein phosphatase regulatory subunit

[Arabidopsis thaliana]

Seq. No. 1362

Contig ID 728_3.R1040



5'-most EST LIB3039-022-Q1-E1-G3

Method BLASTN
NCBI GI g1381675
BLAST score 139
E value 5.0e-72
Match length 303
% identity 87

NCBI Description Glycine max small GTP-binding protein (sra1) mRNA, partial

cds

Seq. No. 1363

Contig ID 728_4.R1040 5'-most EST wrg700786826.h2

Seq. No. 1364

Contig ID 728_5.R1040 5'-most EST gsv701052886.h1

Method BLASTN
NCBI GI g1381675
BLAST score 119
E value 2.0e-60
Match length 267
% identity 86

NCBI Description Glycine max small GTP-binding protein (sra1) mRNA, partial

cds

Seq. No. 1365

Contig ID 728_6.R1040 5'-most EST ncj700988106.h1

Method BLASTN
NCBI GI g1381675
BLAST score 214
E value 1.0e-117
Match length 222
% identity 99

NCBI Description Glycine max small GTP-binding protein (sral) mRNA, partial

cds

Seq. No. 1366

Contig ID 730_1.R1040 5'-most EST awf700841707.h1

Method BLASTX
NCBI GI g1389897
BLAST score 2704
E value 0.0e+00
Match length 643
% identity 86

NCBI Description (U59626) LEA protein [Glycine max]

Seq. No. 1367

Contig ID 732_1.R1040
5'-most EST 94097568_FL
Method BLASTN
NCBI GI 94097568
BLAST score 588
E value 0.0e+00
Match length 619

330



% identity

NCBI Description Glycine max farnesylated protein GMFP4 mRNA, partial cds

Seq. No.

1368

Contig ID

733 4.R1040

5'-most EST

LIB3170-083-Q1-K1-D10

Seq. No.

1369

Contig ID 5'-most EST 733 5.R1040 zpv700761634.h1

Method

BLASTX g3157949

NCBI GI BLAST score

% identity

337

E value Match length 9.0e-34 89 87

NCBI Description

(AC002131) Similar to glucan endo-1,3-beta-D-glucosidase

precursor gb Z28697 from Nicotiana tabacum. ESTs gb Z18185 and gb AA605362 come from this gene. [Arabidopsis thaliana]

Seq. No. Contig ID 1370 733 6.R1040

5'-most EST

jC-gmle01810001b03a1

Seq. No.

1371

Contig ID

733 7.R1040

5'-most EST

uC-qmrominsoy283e09b1

Method NCBI GI BLASTX g4467159

BLAST score

145 4.0e-09

E value Match length

47

% identity

55

NCBI Description

(AL035540) hypothetical protein [Arabidopsis thaliana]

Seq. No.

1372

Contig ID 5'-most EST 733 8.R1040 uxk700671427.h1

Method NCBI GI BLASTX g3157934

BLAST score

E value

509

Match length

6.0e-65 180

% identity NCBI Description

71 (AC002131) Similar to hypothetical protein F09E5.8

gb_U37429 from C. elegans. ESTs gb_T42019 and gb_N97000

come from this gene. [Arabidopsis thaliana]

Seq. No.

1373

Contig ID 5'-most EST 733 9.R1040 asn701136839.h1

Method NCBI GI BLAST score BLASTN q18729 58 8.0e-24

62

98

E value Match length % identity

331



NCBI Description Soybean (Glycine max) 18S ribosomal RNA

Seq. No. 1374

Contig ID 733_10.R1040 5'-most EST gsv701055023.h1

Method BLASTN
NCBI GI g455949
BLAST score 85
E value 6.0e-40
Match length 89
% identity 99

NCBI Description 18S rRNA {3' region} [Glycine max=soybeans, var. Pella 86,

leaves, rRNA Partial, 111 nt]

Seq. No. 1375

Contig ID 733_11.R1040 5'-most EST jsh701069821.h1

Seq. No. 1376

Contig ID 733_12.R1040 5'-most EST uxk700672545.h1

Method BLASTX
NCBI GI g4539303
BLAST score 153
E value 1.0e-20
Match length 132
% identity 47

NCBI Description (AL049480) putative protein [Arabidopsis thaliana]

Seq. No. 1377

Contig ID 733_13.R1040

5'-most EST LIB3106-099-Q1-K1-G12

Method BLASTN
NCBI GI g4115533
BLAST score 63
E value 1.0e-26
Match length 161
% identity 88

NCBI Description Vigna mungo UFGlyT mRNA for UDP-glycose:flavonoid

glycosyltransferase, complete cds

Seq. No. 1378

Contig ID 733 14.R1040

5'-most EST LIB3092-040-Q1-K1-D2

Method BLASTN
NCBI GI g2330648
BLAST score 549
E value 0.0e+00
Match length 1076
% identity 88

NCBI Description Pisum sativum mRNA for topoisomerase I

Seq. No. 1379

Contig ID 733_15.R1040 5'-most EST leu701145478.h1

Method BLASTN NCBI GI g455949



```
BLAST score
E value
                   1.0e-44
Match length
                   93
% identity
                   100
NCBI Description
                  18S rRNA {3' region} [Glycine max=soybeans, var. Pella 86,
                  leaves, rRNA Partial, 111 nt]
Seq. No.
                   1380
Contig ID
                  733 16.R1040
5'-most EST
                  LIB3029-002-Q1-B1-B6
Method
                  BLASTN
NCBI GI
                  q3097320
BLAST score
                  87
E value
                  7.0e-41
Match length
                  577
% identity
                  83
NCBI Description
                  Glycine max gene for Bd 30K, complete cds
Seq. No.
                  1381
Contig ID
                  733 17.R1040
5'-most EST
                  kl1701207520.h1
Method
                  BLASTX
NCBI GI
                  g3738230
BLAST score
                  376
E value
                  7.0e-36
Match length
                  141
% identity
                   54
NCBI Description
                   (AB007790) DREB2A [Arabidopsis thaliana]
                  >gi 4126706 dbj BAA36705 (AB016570) DREB2A [Arabidopsis
                  thaliana]
                  1382
Seq. No.
Contig ID
                  733 18.R1040
5'-most EST
                  LIB3049-021-Q1-E1-G11
Method
                  BLASTX
NCBI GI
                  g4455242
BLAST score
                   304
E value
                  1.0e-27
Match length
                   66
% identity
                  86
NCBI Description
                   (AL035523) geranylgeranylated protein ATGP4 [Arabidopsis
                  thaliana]
Seq. No.
                  1383
                  733 19.R1040
Contig ID
5'-most EST
                   6HA-01-Q1-E1-A2
Seq. No.
                  1384
Contig ID
                  733 20.R1040
5'-most EST
                  LIB3039-052-Q1-E1-E3
```

Method BLASTN NCBI GI q455949 BLAST score 92 9.0e-44E value Match length 96 99 % identity

NCBI Description 18S rRNA {3' region} [Glycine max=soybeans, var. Pella 86,



leaves, rRNA Partial, 111 nt]

Seq. No. 1385 Contig ID 733 21.R1040 5'-most EST leu701154457.h1 Method BLASTX q1168548 NCBI GI BLAST score 329 3.0e-30 E value Match length 110 % identity 59 HOMEOBOX-LEUCINE ZIPPER PROTEIN ATHB-7 (HD-ZIP PROTEIN NCBI Description ATHB-7) >gi_629527_pir_S47137 homeotic protein Athb-7 -Arabidopsis thaliana >gi_499164_emb_CAA47425 (X67032) orf [Arabidopsis thaliana] 1386 Seq. No. Contiq ID 733 22.R1040 5'-most EST LIB3170-066-Q1-J1-H1 Seq. No. 1387 Contig ID 733 23.R1040 zhf700960789.h1 5'-most EST Method BLASTX g2959358 NCBI GI BLAST score 351 4.0e-62 E value Match length 128 % identity 91 (X96758) clathrin coat assembly protein AP17 [Zea mays] NCBI Description Seq. No. 1388 733 24.R1040 Contig ID 5'-most EST zvj700605287.h2 BLASTX Method g2642448 NCBI GI BLAST score 438 E value 8.0e-43 Match length 251 % identity 29 (AC002391) hypothetical protein [Arabidopsis thaliana] NCBI Description >gi_3169187 (AC004401) hypothetical protein [Arabidopsis thaliana] 1389 Seq. No. Contig ID 733 25.R1040 kmv700738553.h1 5'-most EST Method BLASTX

Method BLASTX
NCBI GI g4530591
BLAST score 720
E value 1.0e-95

Match length 242 % identity 71

NCBI Description (AF132475) heme oxygenase 1 [Arabidopsis thaliana]

>gi_4530593_gb_AAD22108.1_ (AF132476) heme oxygenase 1

[Arabidopsis thaliana]



```
1390
Seq. No.
                  733 26.R1040
Contig ID
5'-most EST
                  LIB3072-019-Q1-E1-F5
Method
                  BLASTN
                  g343344
NCBI GI
                  298
BLAST score
                  1.0e-167
E value
Match length
                   628
% identity
                   96
                  Soybean chloroplast 16S rRNA (3' end), 18S rRNA (5' end),
NCBI Description
                   Ile-tRNA, and Ala-tRNA genes
Seq. No.
                  1391
                  733 27.R1040
Contig ID
5'-most EST
                  LIB3051-043-Q1-K1-B7
Method
                  BLASTN
                   q166769
NCBI GI
                  182
BLAST score
                   1.0e-97
E value
                   577
Match length
% identity
                   84
NCBI Description A.thaliana heat shock protein 83 mRNA, complete cds
                   1392
Seq. No.
                   733 28.R1040
Contig ID
5'-most EST
                   LIB3107-079-Q1-K1-F4
Method
                   BLASTN
NCBI GI
                   q4097570
BLAST score
                   626
                   0.0e + 00
E value
Match length
                   982
                   97
% identity
NCBI Description Glycine max farnesylated protein GMFP5 mRNA, partial cds
                   1393
Seq. No.
                   733 29.R1040
Contig ID
5'-most EST
                   asn701140463.hl
Method
                   BLASTN
                   g388206
NCBI GI
BLAST score
                   57
E value
                   3.0e-23
Match length
                   177
% identity
                   84
                  Lycopersicon esculentum ubiquitin carrier protein (Ubc)
NCBI Description
                   mRNA, complete cds. >gi_2086269_gb_I38279_I38279 Sequence
                   109 from patent US
                   1394
Seq. No.
                   733 30.R1040
Contig ID
                   gsv701053339.h1
5'-most EST
```

Method BLASTN g4097570 NCBI GI BLAST score 225 E value 1.0e-123 Match length 787 % identity 91

NCBI Description Glycine max farnesylated protein GMFP5 mRNA, partial cds

Seq. No. 1395

Contig ID 733 31.R1040

5'-most EST LIB3094-038-Q1-K1-A5

Method BLASTN
NCBI GI g1326160
BLAST score 66
E value 3.0e-28
Match length 334
% identity 94

NCBI Description Phaseolus vulgaris dehydrin mRNA, complete cds

Seq. No. 1396

Contig ID 733 32.R1040

5'-most EST LIB3073-017-Q1-K1-D5

Method BLASTN
NCBI GI g2104958
BLAST score 144
E value 6.0e-75
Match length 252
% identity 91

NCBI Description Vicia faba immunophilin (FKBP12) mRNA, complete cds

Seq. No. 1397

Contig ID 733 33.R1040

5'-most EST LIB3074-004-Q1-K1-C12

Method BLASTX
NCBI GI g1350989
BLAST score 219
E value 6.0e-56
Match length 205
% identity 67

NCBI Description PROBABLE 40S RIBOSOMAL PROTEIN S3 >gi_1055070 (U39851)

coded for by C. elegans cDNA CEMSE13F; coded for by C. elegans cDNA CEMSE13R; coded for by C. elegans cDNA

CEMSG70F; coded for by C. elegans cDNA CEESL66F; similar

to ribosomal protein S3 [Caenorhabditis elegans]

Seq. No. 1398

Contig ID 733 34.R1040

5'-most EST LIB3109-038-Q1-K1-A6

Method BLASTX
NCBI GI g133940
BLAST score 278
E value 1.0e-41
Match length 179
% identity 63

NCBI Description 40S RIBOSOMAL PROTEIN S3A (S1A) >gi_70851_pir__R3XL3A

ribosomal protein S3a - African clawed frog

>gi_65091_emb_CAA40592_ (X57322) ribosomal protein S1a

[Xenopus laevis]

Seq. No. 1399

Contig ID 733_35.R1040 5'-most EST awf700837545.h1

Method BLASTX NCBI GI g1389897



BLAST score 437 E value 3.0e-42 Match length 738 % identity 29

NCBI Description (U59626) LEA protein [Glycine max]

Seq. No. 1400

Contig ID 733_36.R1040 5'-most EST hyd700729149.h1

Method BLASTX
NCBI GI g1945447
BLAST score 1140
E value 0.0e+00
Match length 661
% identity 63

NCBI Description (U94395) 90-kDa heat shock protein [Sus scrofa]

Seq. No. 1401

Contig ID 733_37.R1040 5'-most EST ncj700983392.h1

Method BLASTX
NCBI GI g417154
BLAST score 794
E value 0.0e+00
Match length 630
% identity 78

NCBI Description HEAT SHOCK PROTEIN 82 >gi_100685_pir__S25541 heat shock

protein 82 - rice (strain Taichung Native One)

>gi_20256_emb_CAA77978_ (Z11920) heat shock protein 82

(HSP82) [Oryza sativa]

Seq. No. 1402

Contig ID 733 38.R1040

5'-most EST LIB3065-019-Q1-N1-G7

Method BLASTX
NCBI GI g547683
BLAST score 998
E value 1.0e-177
Match length 647
% identity 72

NCBI Description HEAT SHOCK COGNATE PROTEIN 80 >gi_170456 (M96549) heat

shock cognate protein 80 [Solanum lycopersicum]
>gi_445601_prf__1909348A heat shock protein hsp80

[Lycopersicon esculentum]

Seq. No. 1403

Contig ID 733_39.R1040

5'-most EST LIB3040-018-Q1-E1-D11

Method BLASTX
NCBI GI g135404
BLAST score 817
E value 1.0e-114
Match length 419
% identity 54

NCBI Description TUBULIN ALPHA-1/ALPHA-2 CHAIN >gi_81291_pir_S04694 tubulin

alpha chain - Volvox carteri f. nagariensis

>gi 22001 emb CAA31326 (X12846) alpha-1 tubulin [Volvox



carteri] >gi_404089 (L24546) alpha-2 tubulin [Volvox
carteri]

Seq. No. 1404

Contig ID 734 1.R1040

5'-most EST LIB3170-037-Q1-K1-H4

Method BLASTX
NCBI GI 94467145
BLAST score 657
E value 7.0e-69
Match length 154
% identity 81

NCBI Description (AL035540) farnesylated protein (ATFP6) [Arabidopsis

thaliana]

Seq. No. 1405

Contig ID 734_2.R1040 5'-most EST awf700836821.h1

Method BLASTN
NCBI GI 94097572
BLAST score 567
E value 0.0e+00
Match length 583
% identity 99

NCBI Description Glycine max farnesylated protein GMFP7 mRNA, partial cds

Seq. No.

Contig ID 734 3.R1040

5'-most EST LIB3051-084-Q1-K1-A8

1406

Method BLASTN
NCBI GI g4097572
BLAST score 67
E value 3.0e-29
Match length 178

% identity 88

NCBI Description Glycine max farnesylated protein GMFP7 mRNA, partial cds

Seq. No. 1407

Contig ID 735_1.R1040 5'-most EST awf700837424.h1

Method BLASTN
NCBI GI g1732555
BLAST score 570
E value 0.0e+00
Match length 610
% identity 98

NCBI Description Glycine max desiccation protective protein LEA5 (Lea5)

mRNA, complete cds

Seq. No. 1408

Contig ID 737 1.R1040

5'-most EST uC-qmflminsoy047e10b1

Method BLASTN
NCBI GI g1763062
BLAST score 955
E value 0.0e+00
Match length 981

338



% identity Glycine max putative transcription factor SCOF-1 (scof-1) NCBI Description mRNA, complete cds Seq. No. 1409 Contig ID 737 2.R1040 5'-most EST LIB3050-022-Q1-K1-G11 Method BLASTX NCBI GI g1763063 BLAST score 640 E value 1.0e-66 Match length 233 % identity 62 NCBI Description (U68763) SCOF-1 [Glycine max] Seq. No. 1410 737 3.R1040 Contig ID 5'-most EST LIB3092-016-Q1-K1-B12 Method BLASTN NCBI GI q1763062 BLAST score 115 E value 7.0e-58 Match length 281 % identity 93 Glycine max putative transcription factor SCOF-1 (scof-1) NCBI Description mRNA, complete cds Seq. No. 1411 737 4.R1040 Contig ID rlr700899402.h1 5'-most EST Method BLASTX g439493 NCBI GI BLAST score 145 4.0e-09 E value Match length 64 % identity 61 (D26086) zinc-finger protein [Petunia x hybrida] NCBI Description Seq. No. 1412 737 7.R1040 Contig ID 5'-most EST jC-gmle01810050c10a1 1413 Seq. No. 738 1.R1040 Contig ID

5'-most EST

LIB3107-008-Q1-K1-A11 BLASTX

Method NCBI GI BLAST score E value

g2501764 2555

Match length % identity

0.0e + 00490 100

NCBI Description

(U69173) calmodulin-like domain protein kinase isoenzyme

beta [Glycine max]

Seq. No.

1414

Contig ID 5'-most EST 739 4.R1040 crh700854956.h1



Method BLASTN NCBI GI g2501765 BLAST score 136 3.0e-70 E value Match length 364 % identity 84

Glycine max calmodulin-like domain protein kinase isoenzyme NCBI Description

gamma mRNA, complete cds

1415 Seq. No.

739 5.R1040 Contig ID smc700746856.hl 5'-most EST

Method BLASTN NCBI GI g18764 BLAST score 163 E value 2.0e-86 Match length 267 % identity 90

NCBI Description G.max tefS1 gene for elongation factor EF-la

Seq. No.

1416

739 6.R1040 Contig ID

uC-gmrominsoy207f02b1 5'-most EST

Method BLASTN NCBI GI g2501765 BLAST score 287 E value 1.0e-160 Match length 703 % identity 93

Glycine max calmodulin-like domain protein kinase isoenzyme NCBI Description

gamma mRNA, complete cds

1417 Seq. No.

739 7.R1040 Contig ID

uC-gmrominsoy049b08b1 5'-most EST

Method BLASTX g2501766 NCBI GI BLAST score 2789 0.0e + 00E value Match length 538 100 % identity

(U69174) calmodulin-like domain protein kinase isoenzyme NCBI Description

gamma [Glycine max]

1418 Seq. No.

Contig ID 739 8.R1040

LIB3040-013-Q1-E1-C12 5'-most EST

Method BLASTX NCBI GI g1352345 1055 BLAST score E value 0.0e + 00438 Match length % identity

NCBI Description ELONGATION FACTOR 1-ALPHA (EF-1-ALPHA)

>gi 18765 emb_CAA40182 (X56856) eEF-1a [Glycine max]

_3%_1

Seq. No. 1419



```
Contig ID
                   739 9.R1040
5'-most EST
                   crh700853432.h1
Method
                   BLASTN
NCBI GI
                   g2598656
BLAST score
                   600
                   0.0e+00
E value
Match length
                   1360
% identity
                   8.7
NCBI Description Vicia faba mRNA for elongation factor 1-alpha (EF1-a)
Seq. No.
                   1420
                   739_10.R1040
Contig ID
5'-most EST
                   jC-gmro02910011c12a1
Method
                   BLASTN
NCBI GI
                   g2598656
BLAST score
                   662
E value
                   0.0e + 00
                   1510
Match length
% identity
                   88
NCBI Description Vicia faba mRNA for elongation factor 1-alpha (EF1-a)
Seq. No.
                   1421
                   739 11.R1040
Contig ID
5'-most EST
                   jC-gmle01810062e02a1
Method
                   BLASTN
NCBI GI
                   g18764
BLAST score
                   704
E value
                   0.0e + 00
                   1440
Match length
% identity
                   96
NCBI Description G.max tefS1 gene for elongation factor EF-1a
Seq. No.
                   1422
Contig ID
                   740 1.R1040
5'-most EST
                   LIB3029-003-Q1-B1-G2
Method
                   BLASTN
NCBI GI
                   g4097893
BLAST score
                   399
                   0.0e + 00
E value
Match length
                   658
% identity
                   96
NCBI Description
                  Glycine max napin-type 2S albumin 1 precursor, mRNA,
                   complete cds
Seq. No.
                   1423
Contig ID
                   744 1.R1040
```

5'-most EST

zhf700956541.h1

Method BLASTN NCBI GI g1778822 BLAST score 745 0.0e+00E value Match length 757 % identity 100

NCBI Description

Glycine max early light-induced protein (ELIP) mRNA,

complete cds

1424

Seq. No.



```
744 2.R1040
Contig ID
5'-most EST
                   uC-gmropic014a06b1
Method
                   BLASTX
NCBI GI
                   g3650037
BLAST score
                   1336
E value
                   1.0e-180
                   398
Match length
% identity
                   80
NCBI Description
```

(AC005396) auxin-responsive GH3-like protein [Arabidopsis thaliana]

1425 Seq. No. 744 3.R1040 Contig ID rlr700902458.h1 5'-most EST Method BLASTN

NCBI GI q1778822 BLAST score 377 0.0e+00E value 449 Match length 96 % identity

Glycine max early light-induced protein (ELIP) mRNA, NCBI Description

complete cds

1426 Seq. No.

744 4.R1040 Contig ID

LIB3138-081-P1-N1-C3 5'-most EST

Method BLASTN NCBI GI q1778822 BLAST score 230 E value 1.0e-126 Match length 340 % identity 92

Glycine max early light-induced protein (ELIP) mRNA, NCBI Description

complete cds

1427 Seq. No.

Contig ID 745 1.R1040

uC-gmflminsoy054b09b1 5'-most EST

Method BLASTN NCBI GI g2769671 BLAST score 1289 E value 0.0e + 001297 Match length 100 % identity

Glycine max alternative oxidase (Aox2) mRNA, nuclear gene NCBI Description

encoding mitochondrial protein, complete cds

1428 Seq. No.

746_1.R1040 Contig ID 5'-most EST g1946337_FL Method BLASTN q1946337 NCBI GI 1206 BLAST score 0.0e+00E value Match length 1220 % identity

NCBI Description Glycine max alternative oxidase (Aox3) mRNA, nuclear gene



encoding mitochondrial protein, complete cds

Seq. No. 1429
Contig ID 746_2.R1040
5'-most EST wrg700788036.h1

Method BLASTN
NCBI GI g1946337
BLAST score 164
E value 5.0e-87
Match length 253
% identity 84

NCBI Description Glycine max alternative oxidase (Aox3) mRNA, nuclear gene

encoding mitochondrial protein, complete cds

Seq. No. 1430

Contig ID 748 1.R1040

5'-most EST LIB3170-071-Q1-K1-B9

Method BLASTN
NCBI GI g18737
BLAST score 543
E value 0.0e+00
Match length 890
% identity 14

NCBI Description Glycine max RPRP3 gene for repetitive proline-rich protein

Seq. No. 1431

Contig ID 748 2.R1040

5'-most EST uC-gmrominsoy088g05b1

Method BLASTX
NCBI GI g4263704
BLAST score 744
E value 9.0e-79
Match length 246
% identity 61

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1432

748 3.R1040 Contig ID 5'-most EST g4289852 Method BLASTX q4263704 NCBI GI 205 BLAST score E value 1.0e-15 Match length 79 % identity 54

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1433

Contig ID 748 4.R1040

5'-most EST jC-gmst02400023b02a1

Method BLASTX
NCBI GI g4263704
BLAST score 501
E value 2.0e-50
Match length 167



% identity 61

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1434

Contig ID 748_5.R1040

5'-most EST uC-gmropic019e04b1

Method BLASTX
NCBI GI g4263704
BLAST score 495
E value 8.0e-50
Match length 187
% identity 56

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1435

Contig ID 748 6.R1040 5'-most EST q4405568 Method BLASTX NCBI GI q4263704 BLAST score 194 E value 2.0e-28 Match length 128 % identity 59

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1436

Contig ID 748_11.R1040

5'-most EST LIB3139-064-P1-N1-C7

Method BLASTN
NCBI GI g255407
BLAST score 244
E value 1.0e-134
Match length 706
% identity 16

NCBI Description repetitive proline-rich protein {cDNA clone 1A10-2}

[Glycine max=soybeans, axes germinated for 31 hours, mRNA,

943 nt]

Seq. No. 1437

Contig ID 748_13.R1040 5'-most EST kl1701202974.h1

Method BLASTN
NCBI GI g170048
BLAST score 281
E value 1.0e-157
Match length 400
% identity 9

NCBI Description Glycine max SbPRP1 gene encoding a proline-rich protein,

complete cds

Seq. No. 1438

Contig ID 748 16.R1040

5'-most EST LIB3051-011-Q1-E1-F1

Method BLASTN



```
NCBI GI g170065
BLAST score 350
E value 0.0e+00
Match length 825
% identity 26
```

NCBI Description Soybean (G.max) proline-rich cell wall protein (SbPRP2)

gene, complete cds

Seq. No. 1439

Contig ID 751 1.R1040

5'-most EST LIB3139-075-P1-N1-D6

Method BLASTN
NCBI GI g1196896
BLAST score 1168
E value 0.0e+00
Match length 1243
% identity 99

NCBI Description Glycine max acidic ribosomal protein PO mRNA, complete cds

Seq. No. 1440

Contig ID 751 2.R1040

5'-most EST uC-gmropic0001e04b1

Method BLASTN
NCBI GI g1196896
BLAST score 491
E value 0.0e+00
Match length 848
% identity 92

NCBI Description Glycine max acidic ribosomal protein PO mRNA, complete cds

Seq. No. 1441

Contig ID 751_3.R1040 5'-most EST kl1701212451.h1

Method BLASTN
NCBI GI g1196896
BLAST score 227

E value 1.0e-124
Match length 358
% identity 95

NCBI Description Glycine max acidic ribosomal protein PO mRNA, complete cds

Seq. No. 1442

Contig ID 753_1.R1040 5'-most EST ncj700987283.h1

Method BLASTX
NCBI GI g1346501
BLAST score 1517
E value 0.0e+00
Match length 352
% identity 95

NCBI Description MALATE DEHYDROGENASE, GLYOXYSOMAL PRECURSOR

Seq. No. 1443

Contig ID 754_1.R1040 5'-most EST g310569_FL Method BLASTN NCBI GI g310569



```
BLAST score
                  0.0e + 00
E value
                  1004
Match length
                   98
% identity
                  Glycine max seed maturation protein (GmPM3) mRNA, complete
NCBI Description
                  1444
Seq. No.
                  755 1.R1040
Contig ID
                  LIB3049-009-Q1-E1-F10
5'-most EST
Method
                  BLASTX
                   g1170878
NCBI GI
BLAST score
                  2971
                   0.0e + 00
E value
                   563
Match length
% identity
                   100
                  MALATE SYNTHASE, GLYOXYSOMAL (MS) >gi_170026 (L01629)
NCBI Description
                  malate synthase [Glycine max]
                   1445
Seq. No.
                   755 2.R1040
Contig ID
5'-most EST
                   fC-qmst700653114f4
                   BLASTX
Method
                   g2708743
NCBI GI
BLAST score
                   180
                   8.0e-13
E value
Match length
                   184
% identity
                   28
                   (AC003952) putative Tal-1-like reverse transcriptase
NCBI Description
                   [Arabidopsis thaliana]
                   1446
Seq. No.
Contig ID
                   755 3.R1040
                   fC-qmst700653647r1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4262236
BLAST score
                   229
                   2.0e-18
E value
Match length
                   238
% identity
                   30
                   (AC006200) putative ribose 5-phosphate isomerase
NCBI Description
                   [Arabidopsis thaliana]
```

1447 Seq. No.

755 4.R1040 Contig ID

jC-gmst02400054a05a1 5'-most EST

BLASTN Method g170025 NCBI GI 165 BLAST score 1.0e-87 E value 257 Match length % identity 91

NCBI Description Soybean malate synthase (MS) mRNA, 3' end

1448 Seq. No.

Contig ID 755 5.R1040 5'-most EST zpv700758222.hl



```
BLASTN
Method
                  g170025
NCBI GI
BLAST score
                  88
E value
                  7.0e-42
Match length
                  136
% identity
                  91
                  Soybean malate synthase (MS) mRNA, 3' end
NCBI Description
                  1449
Seq. No.
Contig ID
                  755 6.R1040
```

5'-most EST uC-gmflminsoy010h05b1

1450

BLASTN Method g170025 NCBI GI BLAST score 254 1.0e-141 E value Match length 410 % identity 91

NCBI Description Soybean malate synthase (MS) mRNA, 3' end

755 8.R1040 Contig ID 5'-most EST LIB3170-065-Q1-K1-E5 Method BLASTN NCBI GI g170025 BLAST score 185 E value 2.0e-99

Match length 188 99 % identity

Seq. No.

NCBI Description Soybean malate synthase (MS) mRNA, 3' end

Seq. No. 1451 Contig ID 755 9.R1040

txt700736239.h1 5'-most EST

Method BLASTN NCBI GI g170025 BLAST score 407 E value 0.0e+00Match length 575 93 % identity

Soybean malate synthase (MS) mRNA, 3' end NCBI Description

Seq. No. 1452

755 10.R1040 Contig ID

jC-gmst02400008b12d1 5'-most EST

BLASTN Method g170025 NCBI GI BLAST score 75 5.0e-34 E value Match length 107 93 % identity

NCBI Description Soybean malate synthase (MS) mRNA, 3' end

1453 Seq. No.

756_1.R1040 Contig ID jsh701070153.h1 5'-most EST

Method BLASTX NCBI GI g1362051



```
BLAST score
                  0.0e + 00
E value
                  351
Match length
                  94
% identity
                  protein kinase 3 - soybean >gi_310582 (L19361) protein
NCBI Description
                  kinase 3 [Glycine max]
Seq. No.
                  1454
                  756 2.R1040
Contig ID
                  ncj700982236.hl
5'-most EST
                  BLASTX
Method
                  g1161254
NCBI GI
BLAST score
                  1710
                  0.0e+00
E value
                  349
Match length
                  94
% identity
NCBI Description (L38855) protein kinase [Glycine max]
                  1455
Seq. No.
                  756 3.R1040
Contig ID
                  sat701011327.hl
5'-most EST
                  1456
Seq. No.
                  756 4.R1040
Contig ID
                  LIB3138-101-Q1-N1-G4
5'-most EST
                  1457
Seq. No.
                   756 8.R1040
Contig ID
                   eep700869793.hl
5'-most EST
Method
                   BLASTN
                   q1161253
NCBI GI
BLAST score
                   137
E value
                   8.0e-71
                   137
Match length
                   100
% identity
                  Glycine max protein kinase (SPK-4) mRNA, complete cds
NCBI Description
Seq. No.
                   1458
Contig ID
                   756 9.R1040
                   LIB3138-124-Q1-N1-E11
5'-most EST
Method
                   BLASTN
NCBI GI
                   g1161253
BLAST score
                   255
                   1.0e-141
E value
Match length
                   295
% identity
                   97
NCBI Description Glycine max protein kinase (SPK-4) mRNA, complete cds
                   1459
Seq. No.
                   758 1.R1040
Contig ID
                   LIB3109-056-Q1-K1-A4
5'-most EST
                   BLASTN
Method
```

Method BLASTN
NCBI GI g986966
BLAST score 1200
E value 0.0e+00
Match length 1265
% identity 99



NCBI Description Glycine max TGACG-motif binding factor (STF2) mRNA, complete cds

Seq. No. 1460

Contig ID 758 5'-most EST fde

758_2.R1040 fde700873235.h1

Method BLASTN
NCBI GI g2934883
BLAST score 1215
E value 0.0e+00
Match length 1215

% identity 100
NCBI Description Glycine max TGACG-motif binding protein (STF1) mRNA,

complete cds

1461

1462

Seq. No.

Contig ID 5'-most EST 758_5.R1040 hyd700730858.h1

Method BLASTN
NCBI GI g986966
BLAST score 151
E value 2.0e-79
Match length 155
% identity 99

NCBI Description Glycine max TGACG-motif binding factor (STF2) mRNA,

complete cds

Seq. No.

Contig ID 5'-most EST 760_1.R1040 zpv700761071.h1

Method BLASTX
NCBI GI g2353333
BLAST score 211
E value 3.0e-16
Match length 158
% identity 3

NCBI Description (AF016513) Ce-LEA [Caenorhabditis elegans]

>gi_3924819_emb_CAB05543_ (Z83113) cDNA EST EMBL:D26945 comes from this gene; cDNA EST EMBL:D26944 comes from this gene; cDNA EST EMBL:D26943 comes from this gene; cDNA EST EMBL:D26942 comes from this gene; cDNA EST EMBL:D35244

comes from this gene; cDN

Seq. No. 1463

Contig ID 761 1.R1040

5'-most EST uC-gmropic112b10b1

Method BLASTX
NCBI GI g3164115
BLAST score 256
E value 1.0e-21
Match length 145
% identity 38

NCBI Description (AJ224145) major latex-like protein [Rubus idaeus]

Seq. No.

Contig ID 761_2.R1040

5'-most EST LIB3139-048-P1-N1-G3

1464



```
BLASTX
Method
                  q2465015
NCBI GI
BLAST score
                  273
                   1.0e-23
E value
Match length
                  152
% identity
                   40
                  (AJ001449) ripening-induced protein [Fragaria vesca]
NCBI Description
Seq. No.
                   761 3.R1040
Contig ID
5'-most EST
                  LIB3106-006-Q1-K2-C4
                   1466
Seq. No.
                   761 4.R1040
Contig ID
5'-most EST
                   LIB3139-079-P1-N1-B2
Method
                   BLASTX
NCBI GI
                   g2465015
BLAST score
                   267
                   6.0e-23
E value
                   152
Match length
% identity
                   (AJ001449) ripening-induced protein [Fragaria vesca]
NCBI Description
Seq. No.
                   761 5.R1040
Contig ID
5'-most EST
                   asn701134581.h1
                   1468
Seq. No.
                   764 1.R1040
Contig ID
                   LIB3093-045-Q1-K1-H7
5'-most EST
Method
                   BLASTN
                   g21028
NCBI GI
                   469
BLAST score
                   0.0e + 00
E value
                   949
Match length
                   87
% identity
                   P.vulgaris pgip gene for polygalacturonase-inhibiting
NCBI Description
                   protein
                   1469
Seq. No.
                   765 1.R1040
Contig ID
5'-most EST
                   ncj700979033.h1
Method
                   BLASTX
                   g3747089
NCBI GI
                   2160
BLAST score
                   0.0e+00
E value
                   409
Match length
                   100
% identity
                   (AF095445) NADP-dependent isocitrate dehydrogenase [Glycine
NCBI Description
                   max]
```

Seq. No. 1470

Contig ID 765_2.R1040 5'-most EST fua701040651.h1

Method BLASTN
NCBI GI g3747088
BLAST score 320



```
1.0e-180
E value
                  532
Match length
% identity
                  98
                  Glycine max NADP-dependent isocitrate dehydrogenase (IDH1)
NCBI Description
                  mRNA, complete cds
                  1471
Seq. No.
                  765 3.R1040
Contig ID
                  kl1701213049.hl
5'-most EST
                  BLASTN
Method
NCBI GI
                   q3747088
                   91
BLAST score
                   1.0e-43
E value
                   123
Match length
                   93
% identity
                  Glycine max NADP-dependent isocitrate dehydrogenase (IDH1)
NCBI Description
                  mRNA, complete cds
                   1472
Seq. No.
                   769 1.R1040
Contig ID
                   LIB3109-016-Q1-K1-F8
5'-most EST
Method
                   BLASTX
                   g3128220
NCBI GI
                   1167
BLAST score
                   1.0e-128
E value
                   253
Match length
% identity
                   89
                   (AC004077) putative urease accessory protein [Arabidopsis
NCBI Description
                   thaliana] >gi_3337375 (AC004481) putative urease accessory
                   protein [Arabidopsis thaliana]
                   1473
Seq. No.
                   770 1.R1040
Contig ID
                   sat701011252.hl
5'-most EST
                   BLASTX
Method
                   g2462753
NCBI GI
                   1390
BLAST score
                   1.0e-154
E value
                   407
Match length
                   62
% identity
                   (AC002292) putative polygalacturonase [Arabidopsis
NCBI Description
                   thaliana]
                   1474
Seq. No.
                   773 1.R1040
Contig ID
                   LIB3139-032-P1-N1-A4
5'-most EST
                   BLASTX
Method
                   q4559342
NCBI GI
                   3226
BLAST score
                   0.0e + 00
E value
Match length
                   701
 % identity
```

Seq. No. 1475

NCBI Description

Contig ID 774_1.R1040

thaliana]

(AC007087) putative copper methylamine oxidase [Arabidopsis



```
LIB3039-044-Q1-E1-C5
5'-most EST
Method
                  BLASTX
                  g3023858
NCBI GI
BLAST score
                  1690
                  0.0e + 00
E value
                  325
Match length
                  98
% identity
                  GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE
NCBI Description
                  PROTEIN >gi_1256608 (U44850) G protein beta subunit
                  [Glycine max]
                  1476
Seq. No.
                  774 3.R1040
Contig ID
                  LIB3138-087-P1-N1-A5
5'-most EST
Method
                  BLASTN
                  g1256607
NCBI GI
                  270
BLAST score
                  1.0e-150
E value
                  290
Match length
                  98
% identity
NCBI Description Glycine max G protein beta subunit mRNA, complete cds
                  1477
Seq. No.
                  774 4.R1040
Contig ID
5'-most EST
                  LIB3087-012-Q1-K1-B5
                  BLASTN
Method
                  g1256607
NCBI GI
                  166
BLAST score
                  2.0e-88
E value
Match length
                  210
                  96
% identity
NCBI Description Glycine max G protein beta subunit mRNA, complete cds
                  1478
Seq. No.
                  774 5.R1040
Contig ID
5'-most EST
                  LIB3170-039-Q1-K2-G12
Method
                  BLASTX
                  q3023847
NCBI GI
BLAST score
                   144
E value
                   3.0e-09
Match length
                   70
                   57
% identity
                  GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE
NCBI Description
                   PROTEIN >gi_2385376_emb_CAA69934_ (Y08678) G protein beta
                   subunit-like [Medicago sativa]
                   1479
Seq. No.
```

775 1.R1040 Contig ID kl1701203226.h1 5'-most EST Method BLASTN g20562 NCBI GI 152 BLAST score

7.0e-80 E value 368 Match length % identity 85

NCBI Description P.hybrida myb.Ph3 gene encoding protein

Match length

% identity

502

46



```
1480
Seq. No.
                  776 1.R1040
Contig ID
5'-most EST
                  smc700744417.hl
Method
                  BLASTX
NCBI GI
                  g1002800
BLAST score
                  586
                  1.0e-60
E value
Match length
                  126
% identity
                  81
                  (U33917) Cpm7 [Craterostigma plantagineum]
NCBI Description
Seq. No.
                  1481
                  777 1.R1040
Contig ID
5'-most EST
                  gsv701044385.h1
Method
                  BLASTX
NCBI GI
                  g2605621
BLAST score
                  402
                  2.0e-39
E value
                  100
Match length
% identity
                  74
NCBI Description (D88619) OSMYB3 [Oryza sativa]
                  1482
Seq. No.
                  779 1.R1040
Contig ID
5'-most EST
                  LIB3028-029-Q1-B1-E11
Method
                  BLASTX
                  g2129563
NCBI GI
BLAST score
                  466
                   1.0e-46
E value
                   98
Match length
                   88
% identity
                  cM4 protein - Arabidopsis thaliana >gi 928928_emb_CAA62030_
NCBI Description
                   (X90382) cM4 [Arabidopsis thaliana]
Seq. No.
                   1483
Contig ID
                   780_1.R1040
                   fC-gmse700672319f2
5'-most EST
                   BLASTX
Method
                   q1101770
NCBI GI
BLAST score
                   566
                   3.0e-58
E value
                   150
Match length
% identity
                   69
                  (U39448) MYB-like transcriptional factor MBF1 [Picea
NCBI Description
                  mariana]
                   1484
Seq. No.
                   783 1.R1040
Contig ID
5'-most EST
                   gsv701056550.hl
                   BLASTX
Method
NCBI GI
                   q3334665
BLAST score
                   1162
                   1.0e-127
E value
```

NCBI Description (Y10492) putative cytochrome P450 [Glycine max]

Match length

NCBI Description

% identity

55

67



```
Seq. No.
                  783 2.R1040
Contig ID
5'-most EST
                  hyd700730654.h1
                  BLASTX
Method
NCBI GI
                  g4538911
BLAST score
                  676
E value
                  6.0e-71
                  183
Match length
% identity
                  71
                  (ALO49482) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  1486
Seq. No.
                  785 1.R1040
Contig ID
5'-most EST
                  rca700998785.h1
Method
                  BLASTX
NCBI GI
                  g3318611
BLAST score
                  1632
E value
                  0.0e+00
Match length
                  375
                  86
% identity
                   (AB016063) mitochondrial phosphate transporter [Glycine
NCBI Description
                  1487
Seq. No.
Contig ID
                   785 3.R1040
                   jC-gmro02910075b05a1
5'-most EST
Method
                   BLASTN
NCBI GI
                   q3318610
BLAST score
                   257
E value
                   1.0e-142
Match length
                   378
                   93
% identity
                  Glycine max mRNA for mitochondrial phosphate transporter,
NCBI Description
                   complete cds
                   1488
Seq. No.
Contig ID
                   786 1.R1040
                   rlr700899175.h1
5'-most EST
Method
                   BLASTN
NCBI GI
                   g2765444
BLAST score
                   616
E value
                   0.0e+00
                   730
Match length
                   98
% identity
NCBI Description
                  Glycine max mRNA for ferredoxin thioredoxin reductase
                   1489
Seq. No.
                   787 1.R1040
Contig ID
5'-most EST
                   g4294555
Method
                   BLASTX
NCBI GI
                   g1402878
BLAST score
                   208
E value
                   2.0e-16
```

(X98130) unknown [Arabidopsis thaliana]



```
Seq. No.
                    787 2.R1040
Contig ID
                    LIB3028-056-Q1-B1-A5
5'-most EST
                    1491
Seq. No.
                    787 3.R1040
Contig ID
                    uC-\overline{g}mrominsoy200e09b1
5'-most EST
Method
                    BLASTX
                    g4455208
NCBI GI
                    280
BLAST score
                    1.0e-24
E value
```

81

1492

70 % identity (AL035440) putative protein [Arabidopsis thaliana] NCBI Description

787 4.R1040 Contig ID 5'-most EST uC-gmropic041f08b1 Method BLASTN NCBI GI g436789 BLAST score 49 4.0e-18 E value 108 Match length 86 % identity

Match length

Seq. No.

Tobacco RNA-binding glycine-rich protein (RGP-1a) mRNA NCBI Description

1493 Seq. No.

787 5.R1040 Contig ID

5'-most EST LIB3051-083-Q1-K1-D8

BLASTN Method g2267568 NCBI GI BLAST score 56 E value 3.0e-22 Match length 266 83 % identity

Pelargonium x hortorum glycine-rich RNA binding protein 2 NCBI Description

(GRP2) precursor mRNA, complete cds

1494 Seq. No.

787 6.R1040 Contig ID

LIB3106-033-Q1-K1-E8 5'-most EST

BLASTX Method g4455208 NCBI GI 155 BLAST score 7.0e-10 E value 101 Match length % identity 32

(AL035440) putative protein [Arabidopsis thaliana] NCBI Description

1495 Seq. No.

787 7.R1040 Contig ID vwf700678078.hl 5'-most EST

BLASTX Method g544424 NCBI GI BLAST score 163 E value 7.0e-14 84 Match length



% identity GLYCINE-RICH RNA-BINDING PROTEIN 7 >gi 419755_pir__S30147 NCBI Description glycine-rich protein (clone AtGRP7) - Arabidopsis thaliana >gi 16301_emb_CAA78711_ (Z14987) glycine rich protein [Arabidopsis thaliana] >gi_166837 (L00648) RNA-binding protein [Arabidopsis thaliana] >gi 4567224 gb AAD23639.1 AC007119 5 (AC007119) glycine-rich RNA binding protein 7 [Arabidopsis thaliana] 1496 Seq. No. 787 8.R1040 Contig ID LIB3106-060-Q1-K1-B10 5'-most EST BLASTX g2267567 158

Method NCBI GI BLAST score 4.0e-10 E value 91 Match length 59 % identity

(AF009003) glycine-rich RNA binding protein 1 [Pelargonium NCBI Description x hortorum] >gi_2267569 (AF009004) glycine-rich RNA binding

protein 2 [Pelargonium x hortorum]

1497 Seq. No. 787 9.R1040 Contig ID zhf700956043.h1 5'-most EST BLASTX Method g4455208 NCBI GI BLAST score 488

8.0e-49E value 163 Match length % identity 49

(AL035440) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 1498

788 1.R1040 Contig ID

LIB3030-010-Q1-B1-B10 5'-most EST

BLASTX Method q2244811 NCBI GI BLAST score 680 1.0e-73 E value Match length 209 % identity 76

(Z97336) unnamed protein product [Arabidopsis thaliana] NCBI Description

1499 Seq. No.

788 2.R1040 Contig ID g4260108 5'-most EST Method BLASTN g972916 NCBI GI BLAST score 122 1.0e-61 E value Match length 254 % identity

Arabidopsis thaliana IAA7 (IAA7) gene, complete cds NCBI Description

1500 Seq. No.

788_3.R1040 Contig ID



```
LIB3106-034-Q1-K1-F2
5'-most EST
Method
                  BLASTX
                  q117290
NCBI GI
                  1660
BLAST score
                  0.0e+00
E value
                  392
Match length
% identity
                  CHLOROPLAST TRIOSE PHOSPHATE TRANSLOCATOR PRECURSOR (CTPT)
NCBI Description
                  (P36) (E30) >gi 100071 pir__S23774 triose
                  phosphate/3-phosphoglycerate/phosphate translocator
                  precursor - garden pea >gi_20691_emb_CAA38451_ (X54639)
                  chloroplast import receptor p36 [Pisum sativum]
                  >gi 20853 emb CAA48210 (X68077) phosphate translocator
                  [Pisum sativum] >gi_228551_prf__1805409A phosphate
                  translocator [Pisum sativum]
Seq. No.
                  1501
Contig ID
                  788 5.R1040
5'-most EST
                  asn701136914.h1
Method
                  BLASTX
                  q1363487
NCBI GI
                  326
BLAST score
E value
                  3.0e-30
Match length
                  96
% identity
                  71
                  IAA7 protein - Arabidopsis thaliana >gi 972917 (U18409)
NCBI Description
                  IAA7 [Arabidopsis thaliana]
Seq. No.
                   1502
Contig ID
                   790 1.R1040
5'-most EST
                  LIB3073-002-Q1-K1-F3
                  BLASTX
Method
NCBI GI
                   g1717975
                   279
BLAST score
E value
                   1.0e-24
                   98
Match length
% identity
                   49
                  14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) (UK114
NCBI Description
                   ANTIGEN HOMOLOG) >qi 1177435 emb CAA64670 (X95384) 14.5
                   kDa translational inhibitor protein, p14.5 [Homo sapiens]
                   1503
Seq. No.
                   791 1.R1040
Contig ID
                   kl1701208271.h1
5'-most EST
                   BLASTX
Method
                   g3176673
NCBI GI
                   210
BLAST score
                   2.0e-16
E value
                  76
Match length
                   54
% identity
                   (AC003671) Similar to serine/threonine kinase gb_Y12531
NCBI Description
                   from Brassica oleracea. [Arabidopsis thaliana]
```

Seq. No. 1504

Contig ID 791_5.R1040

5'-most EST LIB3170-028-Q1-K1-B10



```
1505
Seq. No.
                  793 1.R1040
Contig ID
                  uxk700667991.hl
5'-most EST
Method
                  BLASTX
                  g687844
NCBI GI
                  283
BLAST score
                  5.0e-25
E value
                  172
Match length
                  5
% identity
                  (U21320) contains TPR domain-like repeats [Caenorhabditis
NCBI Description
                  elegans]
                  1506
Seq. No.
                  793 2.R1040
Contig ID
5'-most EST
                  LIB3049-001-Q1-E1-C7
                  1507
Seq. No.
                  794 1.R1040
Contig ID
                  LIB3106-104-Q1-K1-G8
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3859570
BLAST score
                   322
                   9.0e-30
E value
                   132
Match length
                   54
% identity
                  (AF098753) unknown [Oryza sativa]
NCBI Description
                   1508
Seq. No.
                   799 1.R1040
Contig ID
5'-most EST
                   ssr700556237.h1
                   1509
Seq. No.
                   800 1.R1040
Contig ID
5'-most EST
                   2DA-01-Q1-B1-A9
                   1510
Seq. No.
                   801 1.R1040
Contig ID
                   kmv700738580.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4006868
                   201
BLAST score
E value
                   3.0e-15
                   144
Match length
                   38
% identity
                  (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
                   1511
Seq. No.
                   802 1.R1040
Contig ID
                   g5677474
5'-most EST
Method
                   BLASTX
                   q4193388
NCBI GI
                   507
BLAST score
                   1.0e-51
E value
```

127 Match length % identity 80

(AF091455) translationally controlled tumor protein [Hevea NCBI Description

brasiliensis]



Seq. No. 1512
Contig ID 803_1.R1040
5'-most EST LIB3107-017-Q1-K1-E1
Method BLASTX
NCBI GI g2623199
BLAST score 787
E value 3.0e-84

% identity 99 NCBI Description (AF030290) protein phosphatase X isoform 2 [Arabidopsis

thaliana]

151

Seq. No. 1513

Match length

Contig ID 803_3.R1040 5'-most EST pxt700946264.h1

Method BLASTX
NCBI GI g2623199
BLAST score 273
E value 2.0e-24
Match length 56
% identity 95

NCBI Description (AF030290) protein phosphatase X isoform 2 [Arabidopsis

thaliana]

Seq. No. 1514

Contig ID 805 1.R1040

5'-most EST jC-gmle01810062a02a1

Method BLASTN
NCBI GI g3128138
BLAST score 59
E value 3.0e-24
Match length 478
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MFO20, complete sequence [Arabidopsis thaliana]

Seq. No. 1515

Contig ID 808 1.R1040

5'-most EST LIB3093-005-Q1-K1-A11

Method BLASTX
NCBI GI g1723562
BLAST score 105
E value 5.0e-12
Match length 102
% identity 42

NCBI Description PROBABLE EUKARYOTIC INITIATION FACTOR C17C9.03

>gi 1314155 emb CAA97349_ (Z73099) probable initiation

factor [Schizosaccharomyces pombe]

Seq. No. 1516

Contig ID 808_2.R1040 5'-most EST 2DA-01-Q1-B1-B6

Method BLASTX
NCBI GI g2895097
BLAST score 294
E value 2.0e-26



1518

1519

Match length 128 47 % identity

(AF012072) eIF4GII [Homo sapiens] NCBI Description

>gi 4503541 ref_NP_003751.1 pEIF4G3 UNKNOWN

Seq. No.

Contig ID 5'-most EST 809 1.R1040 pcp700990419.hl

Seq. No.

Contig ID 5'-most EST 810 1.R1040 2DA-01-Q1-B1-B8

BLASTX Method g3426039 NCBI GI 324 BLAST score 6.0e-30 E value Match length 129 51

NCBI Description (AC005168) unknown protein [Arabidopsis thaliana]

Seq. No. Contig ID

% identity

5'-most EST

811 1.R1040 2DA-01-Q1-B1-B9

BLASTX Method g231551 NCBI GI 158 BLAST score 2.0e-10 E value 124 Match length % identity

ANKYRIN, BRAIN VARIANT 2 (ANKYRIN B) (ANKYRIN, NCBI Description

NONERYTHROID)

Seq. No.

811 2.R1040 Contig ID

5'-most EST

uC-gmrominsoy124b05b1

Seq. No.

1521

1520

813 1.R1040 Contig ID

5'-most EST

LIB3051-065-Q1-K1-D1

BLASTX Method g416873 NCBI GI 260 BLAST score 2.0e-22 E value 96 Match length 53 % identity

3-HYDROXYISOBUTYRATE DEHYDROGENASE PRECURSOR (HIBADH) NCBI Description >gi_111295_pir__A32867 3-hydroxyisobutyrate dehydrogenase

(EC 1.1.1.31) precursor - rat (fragment) >gi_556389 (J04628) 3-hydroxyisobutyrate dehydrogenase [Rattus

norvegicus]

Seq. No.

Contig ID 5'-most EST 1522 815 1.R1040 epx701105596.h1

Method BLASTX NCBI GI g2832300 BLAST score 477



```
9.0e-48
E value
                  118
Match length
% identity
                  (AF044285) adenosine-5'-phosphosulfate-kinase [Catharanthus
NCBI Description
                  roseus]
                  1523
Seq. No.
                  817 1.R1040
Contig ID
                  LIB3040-014-Q1-E1-F1
5'-most EST
                  BLASTX
Method
                  q4467098
NCBI GI
                  192
BLAST score
                   1.0e-14
E value
                   58
Match length
% identity
                   (AL035538) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   818 1.R1040
Contig ID
                   gsv701046553.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3881497
                   176
BLAST score
                   6.0e-12
E value
                   318
Match length
                   24
% identity
                   (Z35604) cDNA EST EMBL: Z14593 comes from this gene; cDNA
NCBI Description
                   EST EMBL: T01764 comes from this gene; cDNA EST EMBL: D37668
                   comes from this gene; cDNA EST EMBL:D73892 comes from this
                   gene; cDNA EST EMBL:D71243 comes from this gene; cDN
                   1525
Seq. No.
                   818 2.R1040
Contig ID
                   uC-gmrominsoy300b04b1
5'-most EST
Seq. No.
                   1526
                   819 5.R1040
Contig ID
                   leu701155872.hl
 5'-most EST
                   1527
 Seq. No.
                   819 6.R1040
 Contig ID
                   g4396041
 5'-most EST
                   BLASTX
 Method
                   g2129577
 NCBI GI
                   270
 BLAST score
                   1.0e-23
 E value
                   64
 Match length
                   80
 % identity
                   DnaJ homolog protein - Arabidopsis thaliana >gi_727357
 NCBI Description
                    (U22340) DnaJ homolog [Arabidopsis thaliana]
```

Seq. No. 1528

Contig ID 819_7.R1040 5'-most EST jC-gmle01810022f03a1

Method BLASTX
NCBI GI g4263704
BLAST score 288



1529

E value 8.0e-26 Match length 102 % identity 54

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No.

Contig ID 819_8.R1040

5'-most EST jC-gmf102220088h01a1

Method BLASTX
NCBI GI g4263704
BLAST score 416
E value 3.0e-40
Match length 257
% identity 39

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No.

Contig ID 819 9.R1040 5'-most EST hrw701060405.h1

1530

Method BLASTX
NCBI GI g4263704
BLAST score 336
E value 8.0e-40
Match length 252
% identity 40

NCBI Description (AC006223) putative sugar starvation-induced protein

[Arabidopsis thaliana]

Seq. No. 1531

Contig ID 819_10.R1040

5'-most EST uC-gmropic067e03b1

Method BLASTX
NCBI GI g1213460
BLAST score 568
E value 4.0e-58
Match length 247
% identity 46

NCBI Description (U03374) C subunit of V-ATPase [Amblyomma americanum]

Seq. No.

1532

Contig ID

819 11.R1040

5'-most EST

LIB3107-070-Q1-K1-A6

Method BLASTX
NCBI GI g461944
BLAST score 712
E value 1.0e-132
Match length 419
% identity 54

NCBI Description DNAJ PROTEIN HOMOLOG (DNAJ-1) >gi 18260 emb CAA47925

(X67695) cs DnaJ-1 [Cucumis sativus]

Seq. No. 1533

Contig ID 819 12.R1040

5'-most EST LIB3139-030-P1-N1-F3

Method BLASTN



```
NCBI GI g609222
BLAST score 297
E value 1.0e-166
Match length 1048
% identity 83
NCBI Description P.sativum mRNA for SAMS-1
Seq. No. 1534
```

Contig ID 819 13.R1040 5'-most EST sat701011172.h1 Method BLASTN NCBI GI g3819098

NCBI GI g3819098
BLAST score 205
E value 1.0e-111
Match length 1266
% identity 81

NCBI Description Cicer arietinum mRNA for copper containing amine oxidase

(DAO)

Seq. No. 1535

Contig ID 819 14.R1040

5'-most EST jC-gmst02400051d01a1

Method BLASTN
NCBI GI 9726031
BLAST score 196
E value 1.0e-105
Match length 1006
% identity 82

NCBI Description Actinidia chinensis S-adenosylmethionine synthetase mRNA,

partial cds

Seq. No. 1536

Contig ID 819 15.R1040

5'-most EST LIB3139-097-P1-N1-D5

Method BLASTX
NCBI GI g1125691
BLAST score 420
E value 1.0e-40
Match length 424
% identity 20

NCBI Description (X94301) DnaJ protein [Solanum tuberosum]

Seq. No. 1537

Contig ID 819_16.R1040

5'-most EST LIB3139-013-P1-N1-G4

Method BLASTN
NCBI GI g1724103
BLAST score 204
E value 1.0e-110
Match length 904
% identity 83

NCBI Description Mesembryanthemum crystallinum methionine

adenosyltransferase mRNA, complete cds

Seq. No. 153

Contig ID 819_17.R1040 5'-most EST zhf700962044.h1



```
BLASTN
Method
NCBI GI
                  q609222
BLAST score
                   266
E value
                  1.0e-147
                  848
Match length
                  86
% identity
NCBI Description P.sativum mRNA for SAMS-1
                  1539
Seq. No.
                  822 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy217d08b1
                  BLASTX
Method
NCBI GI
                   g2462827
BLAST score
                   412
                   3.0e-40
E value
                   143
Match length
                   59
% identity
                  (AF000657) probable thiamin biosynthetic enzyme
NCBI Description
                   [Arabidopsis thaliana]
                   1540
Seq. No.
                   824 1.R1040
Contig ID
5'-most EST
                   seb700651176.h1
Method
                   BLASTX
                   g833835
NCBI GI
                   1395
BLAST score
                   1.0e-155
E value
                   486
Match length
% identity
                   55
                  (U26025) amygdalin hydrolase isoform AH I precursor [Prunus
NCBI Description
                   serotina]
                   1541
Seq. No.
                   824 2.R1040
Contig ID
5'-most EST
                   zsq701122505.h1
Method
                   BLASTX
                   g1155255
NCBI GI
BLAST score
                   140
E value
                   1.0e-08
Match length
                   68
% identity
                   40
                  (U39228) beta-glucosidase [Prunus avium]
NCBI Description
Seq. No.
                   1542
                   824 4.R1040
Contig ID
                   zsg701127573.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q1311386
```

Method BLASTX
NCBI GI g1311386
BLAST score 207
E value 2.0e-23
Match length 73
% identity 67

NCBI Description Cyanogenic Beta-Glucosidase Mol_id: 1; Molecule: Cyanogenic

Beta-Glucosidase; Chain: Null; Ec: 3.2.1.21

Seq. No. 1543

Contig ID 825 1.R1040



5'-most EST epx701108418.hl 1544 Seq. No. 827 1.R1040 Contig ID uC-gmropic028d02b1 5'-most EST 1545 Seq. No. 827 2.R1040 Contig ID zhf700958170.hl 5'-most EST 1546 Seq. No. 828 1.R1040 Contig ID 5'-most EST jex700908654.h1Method BLASTX g4309736 NCBI GI BLAST score 244 3.0e-20 E value 61 Match length 77 % identity (AC006439) hypothetical protein [Arabidopsis thaliana] NCBI Description 1547 Seq. No. 830 1.R1040 Contig ID 2DA-01-Q1-B1-D4 5'-most EST Method BLASTX g2642158 NCBI GI 457 BLAST score E value 2.0e-45 115 Match length % identity 72 NCBI Description (AC003000) hypothetical protein [Arabidopsis thaliana] 1548 Seq. No. Contig ID 830 2.R1040 5'-most EST LIB3106-006-Q1-K2-C1 Method BLASTX NCBI GI q2642158 295 BLAST score 1.0e-26 E value Match length 80 % identity (AC003000) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 1549 832 1.R1040 Contig ID LIB3040-003-Q1-E1-H5 5'-most EST Method BLASTX NCBI GI q417360 BLAST score 369 5.0e-35 E value Match length 133

53 % identity

HIGH MOBILITY GROUP-LIKE NUCLEAR PROTEIN 2 NCBI Description

>gi 2131280_pir__S67767 high mobility group-like protein

NHP2 - yeast (Saccharomyces cerevisiae)

>gi 666101 emb CAA40885 (X57714) high mobility group-like

nuclear protein 2 [Saccharomyces cerevisiae]



>gi 1429348 emb CAA67483 (X99000) high-mobility-group-like protein [Saccharomyces cerevisiae]

>gi 1431346_emb_CAA98786_ (Z74256) ORF YDL208w

[Saccharomyces cerevisiae]

1550 Seq. No.

832 3.R1040 Contig ID ncj700980996.hl 5'-most EST

Method BLASTX g417360 NCBI GI 281 BLAST score 5.0e-25 E value 80 Match length % identity 64

HIGH MOBILITY GROUP-LIKE NUCLEAR PROTEIN 2 NCBI Description

>gi_2131280_pir__S67767 high mobility group-like protein

NHP2 - yeast (Saccharomyces cerevisiae)

>gi 666101 emb CAA40885 (X57714) high mobility group-like

nuclear protein 2 [Saccharomyces cerevisiae]

>gi 1429348 emb_CAA67483_ (X99000) high-mobility-group-like

protein [Saccharomyces cerevisiae]

>gi_1431346_emb_CAA98786 (Z74256) ORF YDL208w

[Saccharomyces cerevisiae]

1551 Seq. No.

832 4.R1040 Contig ID

LIB3072-017-Q1-E1-C5 5'-most EST

BLASTX Method q417360 NCBI GI BLAST score 241 2.0e-20 E value 76 Match length 59 % identity

NCBI Description

HIGH MOBILITY GROUP-LIKE NUCLEAR PROTEIN 2

>gi_2131280_pir__S67767 high mobility group-like protein

NHP2 - yeast (Saccharomyces cerevisiae)

>gi_666101_emb_CAA40885_ (X57714) high mobility group-like

nuclear protein 2 [Saccharomyces cerevisiae]

>gi_1429348_emb_CAA67483_ (X99000) high-mobility-group-like

protein [Saccharomyces cerevisiae]

>gi_1431346_emb_CAA98786_ (Z74256) ORF YDL208w

[Saccharomyces cerevisiae]

1552 Seq. No.

833 1.R1040 Contig ID 2DA-01-Q1-B1-D7 5'-most EST

Method BLASTX NCBI GI q3386604 207 BLAST score 2.0e-16 E value 129 Match length 39 % identity

(AC004665) putative protein kinase [Arabidopsis thaliana] NCBI Description

Seq. No.

1553

834 1.R1040 Contig ID bth700845265.h1 5'-most EST



```
BLASTX
Method
                   q3912919
NCBI GI
BLAST score
                   456
E value
                   1.0e-45
Match length
                   116
                   80
% identity
                   (AF001308) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1554
                   835 1.R1040
Contig ID
                   LIB3138-079-P1-N1-E2
5'-most EST
                   BLASTN
Method
                   g14311
NCBI GI
BLAST score
                   850
E value
                   0.0e + 00
Match length
                   943
                   98
% identity
                   G.max chloroplast gene rps19 and flanking regions
NCBI Description
                   1555
Seq. No.
                   835 2.R1040
Contig ID
                   jC-gmle01810070a03a1
5'-most EST
Method
                   BLASTN
                   g343648
NCBI GI
BLAST score
                   299
E value
                   1.0e-167
                   661
Match length
% identity
                   91
                   Vigna unguiculata chloroplast ribosomal protein (L16 and
NCBI Description
                   L14) genes, 3' and 5' end respectively
                   1556
Seq. No.
                   835 3.R1040
Contig ID
                   LIB3139-038-P1-N1-H6
 5'-most EST
                   BLASTN
Method
                   q11748
NCBI GI
BLAST score
                    524
                    0.0e + 00
E value
                    730
Match length
                    94
 % identity
                   Nicotiana debneyi chloroplast DNA for tRNA-Ile and
NCBI Description
                    ribosomal protein L2
                    1557
 Seq. No.
 Contig ID
                    835 4.R1040
                    ujr700646647.hl
 5'-most EST
                    BLASTN
 Method
                    q12212
 NCBI GI
 BLAST score
                    337
                    0.0e + 00
 E value
 Match length
                    505
```

93 % identity

S.alba chloroplast rp123 and rp12 genes for ribosomal NCBI Description

proteins L23 and L2

1558 Seq. No.

835 9.R1040 Contig ID



LIB3107-071-Q1-K1-H2 5'-most EST BLASTN Method q12212 NCBI GI 141 BLAST score 2.0e-73 E value 185 Match length 94 % identity S.alba chloroplast rp123 and rp12 genes for ribosomal NCBI Description proteins L23 and L2 1559 Seq. No. 835 10.R1040 Contig ID LIB3138-042-Q1-N1-H5 5'-most EST BLASTN Method g11748 NCBI GI 289 BLAST score 1.0e-161 E value 397 Match length 95 % identity Nicotiana debneyi chloroplast DNA for tRNA-Ile and NCBI Description ribosomal protein L2 1560 Seq. No. 835 13.R1040 Contig ID jC-gmle01810032d10d1 5'-most EST BLASTX Method g133939 NCBI GI 314 BLAST score 9.0e-29 E value 80 Match length 72 % identity CHLOROPLAST 30S RIBOSOMAL PROTEIN S3 >gi_70865_pir__R3NT3 NCBI Description ribosomal protein S3 - common tobacco chloroplast >gi_11865_emb_CAA77381_ (Z00044) ribosomal protein S3 [Nicotiana tabacum] >gi_225235_prf__1211235BT ribosomal protein S3 [Nicotiana tabacum] 1561 Seq. No. 836 1.R1040 Contig ID hyd700730304.hl 5'-most EST 1562 Seq. No. 837 1.R1040 Contig ID uC-gmropic018e03b1 5'-most EST Method BLASTX g2244898 NCBI GI BLAST score 186 E value 2.0e-13 96 Match length % identity 44

NCBI Description

(Z97338) strong similarity to protein phosphatase 2A regulatory chain, 74K [Arabidopsis thaliana]

Seq. No. 1563

Contig ID 837_3.R1040 5'-most EST zzp700831795.h1

e was present



Seq. No. 1564

Contig ID 838_1.R1040 5'-most EST ssr700561089.h1

Method BLASTN
NCBI GI g12139
BLAST score 639
E value 0.0e+00
Match length 3513
% identity 90

NCBI Description Pea plastid genes rps2, atpI, atpH, atpF, atpA, trnR and trnG coding for ribosomal protein S2, one CF(1) and three CF(0) subunits of ATP synthase and tRNA-Arg and tRNA-Gly

Seq. No. 1565

Contig ID 838_2.R1040

5'-most EST fC-gmle700553639a4

Method BLASTN
NCBI GI g12139
BLAST score 145
E value 1.0e-75
Match length 414
% identity 91

NCBI Description Pea plastid genes rps2, atpI, atpH, atpF, atpA, trnR and trnG coding for ribosomal protein S2, one CF(1) and three CF(0) subunits of ATP synthase and tRNA-Arg and tRNA-Gly

. No. 1566

Seq. No. 1566
Contig ID 838_4.R1040
5'-most EST zsg701130050.h1

Method BLASTN
NCBI GI g12249
BLAST score 38
E value 5.0e-12
Match length 54
% identity 93

NCBI Description Spinach plastid genes atpI-H-F for ATP synthase CF(O)

subunits IV, III and I

Seq. No. 1567

Contig ID 838 11.R1040

5'-most EST jC-gmle01810075f04a1

Method BLASTN
NCBI GI g12144
BLAST score 41
E value 1.0e-13
Match length 82
% identity 94

NCBI Description Pea chloroplast genes for F(O)-ATP synthase subunits a and

c (partial)

Seq. No. 1568

Contig ID 841_1.R1040 5'-most EST 2DA-01-Q1-B1-E4

Method BLASTX
NCBI GI g2853087
BLAST score 199
E value 3.0e-15



```
122
Match length
                  26
% identity
NCBI Description (AL021768) putative protein [Arabidopsis thaliana]
Seq. No.
                  1569
                  842 1.R1040
Contig ID
                  wvk700685058.hl
5'-most EST
                  BLASTX
Method
                  q3212854
NCBI GI
BLAST score
                  410
                  2.0e-39
E value
                  297
Match length
% identity
NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]
                   1570
Seq. No.
                   842 2.R1040
Contig ID
                   LIB3107-004-Q1-K1-C2
5'-most EST
                   BLASTX
Method
                   g3212854
NCBI GI
BLAST score
                   149
                   2.0e-09
E value
                   57
Match length
                   49
% identity
NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]
                   1571
Seq. No.
                   842 3.R1040
Contig ID
                   zzp700835324.hl
5'-most EST
                   1572
Seq. No.
                   843 1.R1040
Contig ID
                   2DA-01-Q1-B1-E6
5'-most EST
                   BLASTX
Method
                   g3641870
NCBI GI
                   218
BLAST score
                   1.0e-17
E value
                   136
Match length
 % identity
                   41
                   (AJ011013) hypothetical protein [Cicer arietinum]
 NCBI Description
                   1573
 Seq. No.
                   844 1.R1040
 Contig ID
                   uC-gmrominsoy083e10b1
 5'-most EST
                   BLASTX
 Method
                   g2598579
 NCBI GI
                   1461
 BLAST score
                   1.0e-162
 E value
 Match length
                   364
                   74
 % identity
                   (Y15295) L-ascorbate oxidase [Medicago truncatula]
 NCBI Description
```

Seq. No. 844 2.R1040 Contig ID hyd700725411.h1 5'-most EST

1574

Method BLASTN g2598578 NCBI GI

NCBI Description



```
238
BLAST score
                  1.0e-131
E value
Match length
                   690
                   84
% identity
NCBI Description Medicago truncatula mRNA for L-ascorbate oxidase
                   1575
Seq. No.
                   844 4.R1040
Contig ID
                   g4287777
5'-most EST
Method
                   BLASTX
                   g2598579
NCBI GI
BLAST score
                   161
                   3.0e-11
E value
                   49
Match length
% identity
                   (Y15295) L-ascorbate oxidase [Medicago truncatula]
NCBI Description
                   1576
Seq. No.
                   846 1.R1040
Contig ID
                   jex700908385.h1
5'-most EST
Method
                   BLASTX
                   g2462762
NCBI GI
                   961
BLAST score
                   1.0e-104
E value
                   338
Match length
% identity
                   58
                   (AC002292) Highly similar to auxin-induced protein
NCBI Description
                   (aldo/keto reductase family) [Arabidopsis thaliana]
                   1577
Seq. No.
                   846 2.R1040
Contig ID
                   LIB3109-009-Q1-K1-E2
5'-most EST
                   BLASTX
Method
                   g2462762
NCBI GI
BLAST score
                   371
                   3.0e-35
E value
                   141
Match length
                   57
% identity
                   (AC002292) Highly similar to auxin-induced protein
NCBI Description
                   (aldo/keto reductase family) [Arabidopsis thaliana]
                   1578
Seq. No.
                   846 4.R1040
Contig ID
                   LIB3170-068-Q1-J1-D4
5'-most EST
                   1579
Seq. No.
                   846 9.R1040
Contig ID
                   qsv701053511.h1
 5'-most EST
Method
                   BLASTX
                   q2462762
NCBI GI
BLAST score
                   113
                   7.0e-14
E value
                   78
Match length
                    55
 % identity
```

(AC002292) Highly similar to auxin-induced protein

(aldo/keto reductase family) [Arabidopsis thaliana]



Seq. No. 1580

Contig ID 848_1.R1040

5'-most EST LIB3106-033-Q1-K1-C2

Method BLASTN
NCBI GI g4220631
BLAST score 97
E value 7.0e-47
Match length 225
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K5J14, complete sequence [Arabidopsis thaliana]

Seq. No. 1581

Contig ID 848 2.R1040

5'-most EST LIB3056-010-Q1-N1-D9

Method BLASTN
NCBI GI g4220631
BLAST score 93
E value 1.0e-44
Match length 225
% identity 85

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K5J14, complete sequence [Arabidopsis thaliana]

Seq. No. 1582

Contig ID 848_3.R1040

5'-most EST LIB3049-025-Q1-E1-G4

Method BLASTN
NCBI GI g4220631
BLAST score 89
E value 2.0e-42
Match length 225
% identity 85

% identity 85 NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K5J14, complete sequence [Arabidopsis thaliana]

Seq. No. 1583

Contig ID 848_6.R1040 5'-most EST pmv700888989.h1

Method BLASTN
NCBI GI g4220631
BLAST score 65
E value 3.0e-28
Match length 181

% identity 85 NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K5J14, complete sequence [Arabidopsis thaliana]

Seq. No. 1584

Contig ID 849_1.R1040 5'-most EST 2DA-01-Q1-B1-F11

Method BLASTX
NCBI GI g3461848
BLAST score 179
E value 8.0e-13
Match length 102
% identity 49



```
(AC005315) putative ATPase [Arabidopsis thaliana]
NCBI Description
                  1585
Seq. No.
                  850_1.R1040
Contig ID
                  uC-gmropic026d05b1
5'-most EST
                  BLASTX
Method
                   g3688162
NCBI GI
BLAST score
                   654
                   3.0e-68
E value
                   168
Match length
                   79
% identity
                  (AJ009672) centrin [Arabidopsis thaliana]
NCBI Description
                   1586
Seq. No.
                   850 2.R1040
Contig ID
                   ssr700555006.h1
5'-most EST
                   BLASTX
Method
                   g3386622
NCBI GI
BLAST score
                   227
                   2.0e-18
E value
                   129
Match length
% identity
                   65
                   (AC004665) unknown protein [Arabidopsis thaliana]
NCBI Description
                   1587
Seq. No.
                   850_3.R1040
Contig ID
                   2DA-01-Q1-B1-F12
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3688162
                   491
BLAST score
                   2.0e-49
E value
                   135
Match length
                   52
% identity
                   (AJ009672) centrin [Arabidopsis thaliana]
NCBI Description
                   1588
Seq. No.
                   852 1.R1040
Contig ID
                   LIB3138-005-Q1-N1-F7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2315451
BLAST score
                   259
                   8.0e-22
E value
                   142
Match length
                   30
% identity
                    (AF016448) No definition line found [Caenorhabditis
NCBI Description
                   elegans]
                    1589
Seq. No.
 Contig ID
                    853 1.R1040
 5'-most EST
                    2DA-01-Q1-B1-F4
                    1590
 Seq. No.
```

Contig ID 854_1.R1040

5'-most EST LIB3092-017-Q1-K1-D4

Method BLASTX
NCBI GI g2708532
BLAST score 1184



```
1.0e-130
E value
                  344
Match length
% identity
                  67
                  (AF029351) putative RNA binding protein [Nicotiana tabacum]
NCBI Description
                  1591
Seq. No.
                  854 2.R1040
Contig ID
                  q4397165
5'-most EST
                  1592
Seq. No.
                  855 1.R1040
Contig ID
                  LIB3109-030-Q1-K1-B11
5'-most EST
                  1593
Seq. No.
Contig ID
                  856 1.R1040
                  uC-gmflminsoy040a11b1
5'-most EST
                  1594
Seq. No.
                  857 1.R1040
Contig ID
                  bth700845764.h1
5'-most EST
                  1595
Seq. No.
                   858 1.R1040
Contig ID
                   uC-gmropic078g11b1
5'-most EST
                   BLASTX
Method
                   g629692
NCBI GI
                   1174
BLAST score
                   1.0e-129
E value
                   533
Match length
                   48
% identity
                   hypothetical protein - common tobacco
NCBI Description
                   >gi_506471_emb_CAA56189_ (X79794) unnamed protein product
                   [Nicotiana tabacum]
                   1596
Seq. No.
                   858 2.R1040
Contig ID
5'-most EST
                   zhf700955944.hl
                   BLASTX
Method
                   q629692
NCBI GI
BLAST score
                   846
E value
                   5.0e-91
                   184
Match length
% identity
                   hypothetical protein - common tobacco
NCBI Description
                   >gi_506471_emb_CAA56189_ (X79794) unnamed protein product
                   [Nicotiana tabacum]
                   1597
Seq. No.
                   861 1.R1040
Contig ID
 5'-most EST
                   dpv701100030.h2
Method
                   BLASTX
                   g1737492
NCBI GI
                   1936
 BLAST score
                   0.0e+00
E value
Match length
                   624
                   49
 % identity
NCBI Description (U81318) poly(A)-binding protein [Triticum aestivum]
```

Method

NCBI GI



```
1598
Seq. No.
                  861 2.R1040
Contig ID
                  LIB3051-006-Q1-E1-A7
5'-most EST
                  1599
Seq. No.
                  861 4.R1040
Contig ID
                  uC-gmronoir007h06b1
5'-most EST
                  BLASTX
Method
                  g1171978
NCBI GI
                   156
BLAST score
                   2.0e-10
E value
                   48
Match length
% identity
                   67
                  POLYADENYLATE-BINDING PROTEIN 2 (POLY(A) BINDING PROTEIN 2)
NCBI Description
                   (PABP 2) >gi_304109 (L19418) poly(A)-binding protein
                   [Arabidopsis thaliana] >gi_2911051_emb_CAA17561_ (AL021961)
                   poly(A)-binding protein [Arabidopsis thaliana]
                   1600
Seq. No.
                   863_1.R1040
Contig ID
                   2DA-02-Q1-B1-D6
5'-most EST
                   BLASTX
Method
                   g1279224
NCBI GI
BLAST score
                   400
                   2.0e-38
E value
                   216
Match length
% identity
                   (X97454) specific tissue protein 1 [Cicer arietinum]
NCBI Description
                   1601
Seq. No.
Contig ID
                   863 2.R1040
                   LIB3107-034-Q1-K1-B5
5'-most EST
Method
                   BLASTX
                   q3393062
NCBI GI
BLAST score
                   627
E value
                   4.0e-65
                   208
Match length
% identity
                   (Y17386) putative In2.1 protein [Triticum aestivum]
NCBI Description
                   1602
Seq. No.
                   863 4.R1040
Contig ID
5'-most EST
                   uC-gmropic070d05b1
Method
                   BLASTX
                   g1279224
NCBI GI
BLAST score
                   324
E value
                   7.0e-30
                   186
Match length
% identity
                   32
                   (X97454) specific tissue protein 1 [Cicer arietinum]
NCBI Description
                   1603
Seq. No.
                   864 1.R1040
 Contig ID
                   g5752500
 5'-most EST
                   BLASTX
```

g3328240



```
966
BLAST score
                  1.0e-105
E value
                  289
Match length
% identity
                  66
                  (AF064775) early nodule-specific protein [Medicago
NCBI Description
                  truncatula]
                  1604
Seq. No.
                  865 1.R1040
Contig ID
                   zhf700952383.hl
5'-most EST
                  BLASTX
Method
                   g4210330
NCBI GI
                   2927
BLAST score
                   0.0e+00
E value
                   643
Match length
                   83
% identity
                   (AJ223802) 2-oxoglutarate dehydrogenase, El subunit
NCBI Description
                   [Arabidopsis thaliana]
                   1605
Seq. No.
                   865 3.R1040
Contig ID
                   zzp700832947.hl
5'-most EST
                   1606
Seq. No.
                   866_1.R1040
Contig ID
                   2DA-01-Q1-B1-G6
5'-most EST
                   1607
Seq. No.
                   868 1.R1040
Contig ID
                   2DA-01-Q1-B1-G8
5'-most EST
                   BLASTX
Method
                   g1402878
NCBI GI
                   711
BLAST score
                   4.0e-75
E value
                   185
Match length
% identity
                   (X98130) unknown [Arabidopsis thaliana]
NCBI Description
                   1608
Seq. No.
                   869 1.R1040
Contiq ID
                   LIB3170-064-Q1-K1-A3
5'-most EST
                   1609
Seq. No.
                   869 2.R1040
Contig ID
                   bth700845877.h1
5'-most EST
                   BLASTX
Method
                   g3879119
NCBI GI
BLAST score
                   345
                   4.0e-32
E value
Match length
                   168
                   48
 % identity
                    (Z70310) similar to Glutathione S-transferases.
```

NCBI Description

Seq. No. 1610 Contig ID 871_1.R1040 5'-most EST 2DA-01-Q1-B1-H10

[Caenorhabditis elegans]

5'-most EST

Method



```
BLASTX
Method
                  q4512703
NCBI GI
                  146
BLAST score
                  7.0e-14
E value
Match length
                  117
% identity
                  38
NCBI Description (AC006569) unknown protein [Arabidopsis thaliana]
                   1611
Seq. No.
                  871 2.R1040
Contig ID
                  LIB3050-001-Q1-E1-D12
5'-most EST
                   1612
Seq. No.
                   871 3.R1040
Contig ID
                   LIB3051-069-Q1-K1-A10
5'-most EST
Seq. No.
                   1613
                   872 1.R1040
Contig ID
                   LIB3093-050-Q1-K1-C9
5'-most EST
                   BLASTX
Method
                   q3047104
NCBI GI
BLAST score
                   1379
                   1.0e-163
E value
                   469
Match length
% identity
                   (AF058919) No definition line found [Arabidopsis thaliana]
NCBI Description
                   1614
Seq. No.
                   872 2.R1040
Contig ID
                   fC-gmf1700905215y1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3047104
                   419
BLAST score
                   1.0e-40
E value
                   202
Match length
                   61
% identity
                   (AF058919) No definition line found [Arabidopsis thaliana]
NCBI Description
                   1615
Seq. No.
                   873 2.R1040
Contig ID
                   LIB3138-023-Q1-N1-C5
5'-most EST
                   BLASTX
Method
                   g4510383
NCBI GI
                   214
BLAST score
E value
                   4.0e-17
Match length
                   67
                   66
% identity
                   (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
                   1616
Seq. No.
                   873 3.R1040
 Contig ID
                   2DA-01-Q1-B1-H4
 5'-most EST
                   1617
 Seq. No.
                   873 4.R1040
 Contig ID
```

377

LIB3139-059-P1-N1-E8

BLASTX

5'-most EST

Method

BLASTX



```
q4510383
NCBI GI
                  188
BLAST score
                  4.0e-14
E value
Match length
                  63
% identity
                  (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
                  1618
Seq. No.
                  874 1.R1040
Contig ID
                  pxt700945347.hl
5'-most EST
                  BLASTX
Method
                   g4539292
NCBI GI
                   465
BLAST score
                   4.0e-46
E value
Match length
                   96
                   89
% identity
                   (AL049480) putative ribosomal protein S10 [Arabidopsis
NCBI Description
                   thaliana]
                   1619
Seq. No.
                   878 1.R1040
Contig ID
                   vwf700677625.h1
5'-most EST
                   BLASTX
Method
                   g2980806
NCBI GI
                   338
BLAST score
                   8.0e-42
E value
                   207
Match length
                   35
% identity
                  (AL022197) putative protein [Arabidopsis thaliana]
NCBI Description
                   1620
Seq. No.
                   878 3.R1040
Contig ID
                   sat701013749.h1
5'-most EST
                   1621
Seq. No.
                   878 4.R1040
Contig ID
5'-most EST
                   xpa700792783.hl
                   1622
Seq. No.
Contig ID
                   880 1.R1040
                   2DA-01-Q1-B1-H9
5'-most EST
                   1623
Seq. No.
                   881 1.R1040
Contig ID
                   zsg701120314.h1
 5'-most EST
                   BLASTX
Method
                   g4454032
NCBI GI
 BLAST score
                   712
                   5.0e-75
 E value
 Match length
                   243
 % identity
                   60
                   (AL035394) putative protein [Arabidopsis thaliana]
 NCBI Description
                   1624
 Seq. No.
                   883 1.R1040
 Contig ID
                   asj700967313.hl
```

378

```
g3953471
NCBI GI
BLAST score
                  261
                  2.0e-22
E value
                  100
Match length
% identity
                  55
                  (AC002328) F2202.16 [Arabidopsis thaliana]
NCBI Description
                  1625
Seq. No.
                  883 2.R1040
Contig ID
                  fde700873705.h1
5'-most EST
                  1626
Seq. No.
                  885_1.R1040
Contig ID
5'-most EST
                  wvk700682626.h2
                  1627
Seq. No.
                  886 1.R1040
Contig ID
5'-most EST
                  q5688325
Method
                  BLASTX
NCBI GI
                  q1172584
BLAST score
                  1389
                  1.0e-154
E value
                  428
Match length
                  63
% identity
NCBI Description
                  POLYPHENOL OXIDASE PRECURSOR (PPO) (CATECHOL OXIDASE)
                  >gi_1076478_pir__S52984 catechol oxidase (EC 1.10.3.1)
                  precursor - apple tree >gi 507280 (L29450) polyphenol
                  oxidase [Malus domestica]
Seq. No.
                  1628
                  886 2.R1040
Contig ID
                  LIB3106-111-Q1-K1-G4
5'-most EST
                  BLASTX
Method
NCBI GI
                  q1172584
BLAST score
                  509
E value
                  2.0e-51
Match length
                  207
                   49
% identity
                  POLYPHENOL OXIDASE PRECURSOR (PPO) (CATECHOL OXIDASE)
NCBI Description
                  >gi_1076478_pir__S52984 catechol oxidase (EC 1.10.3.1)
                  precursor - apple tree >gi 507280 (L29450) polyphenol
                  oxidase [Malus domestica]
Seq. No.
                  1629
                  887 1.R1040
Contig ID
                  taw700660372.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q728827
BLAST score
                  612
```

Method BLASTX
NCBI GI g728827
BLAST score 612
E value 4.0e-63
Match length 322
% identity 40

NCBI Description ALLANTOINASE PRECURSOR >gi_458126 (U03471) allantoinase

[Rana catesbeiana]

Seq. No. 1630

Contig ID 887_2.R1040

NCBI GI

BLAST score

260



```
smc700744903.h1
5'-most EST
                  BLASTX
Method
                  g728827
NCBI GI
                  185
BLAST score
                  1.0e-13
E value
                  93
Match length
                  43
% identity
                  ALLANTOINASE PRECURSOR >gi_458126 (U03471) allantoinase
NCBI Description
                  [Rana catesbeiana]
                  1631
Seq. No.
                  887 3.R1040
Contig ID
                  hrw701061511.hl
5'-most EST
Seq. No.
                  1632
                  889 1.R1040
Contig ID
                  LIB3107-056-Q1-K1-C3
5'-most EST
                  BLASTX
Method
                  g1652704
NCBI GI
                  186
BLAST score
                  2.0e-13
E value
                  95
Match length
                  43
% identity
NCBI Description (D90907) hypothetical protein [Synechocystis sp.]
                  1633
Seq. No.
                  890 1.R1040
Contig ID
                  gsv701054306.hl
5'-most EST
                  BLASTX
Method
                  g4337195
NCBI GI
                   1386
BLAST score
                   1.0e-154
E value
                   377
Match length
                   78
% identity
NCBI Description (AC006403) putative protein kinase [Arabidopsis thaliana]
                   1634
Seq. No.
                   890 2.R1040
Contig ID
                   zsq701125396.hl
5'-most EST
Method
                   BLASTX
                   q4337195
NCBI GI
                   537
BLAST score
                   8.0e-55
E value
                   120
Match length
% identity
                   (AC006403) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   1635
Seq. No.
                   890 3.R1040
Contig ID
                   uC-gmrominsoy225b07b1
5'-most EST
Seq. No.
                   1636
                   890 5.R1040
Contig ID
                   gsv701055170.hl
5'-most EST
Method
                   BLASTX
                   q4337195
```



```
2.0e-22
E value
                  60
Match length
                  77
% identity
                  (AC006403) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                  1637
Seq. No.
                  893 1.R1040
Contig ID
                  LIB3039-043-Q1-E1-C8
5'-most EST
                  BLASTN
Method
                  g1902893
NCBI GI
                  231
BLAST score
                  1.0e-126
E value
                  603
Match length
% identity
                  85
NCBI Description Solanum melongena mRNA for QM family protein, complete cds
                  1638
Seq. No.
                   893 2.R1040
Contig ID
                  wvk700686407.h1
5'-most EST
                  BLASTX
Method
                   g1169494
NCBI GI
                   1343
BLAST score
                   0.0e+00
E value
                   479
Match length
% identity
                   87
                   ELONGATION FACTOR TU, CHLOROPLAST PRECURSOR (EF-TU)
NCBI Description
                   >gi 2119915_pir__S60659 EF-Tu protein precursor - soybean
                   >gi_949873_emb_CAA61444_ (X89058) EF-Tu protein [Glycine
                   max]
                   1639
Seq. No.
                   893 3.R1040
Contig ID
                   LIB3106-106-Q1-K1-A5
5'-most EST
Method
                   BLASTN
                   q18775
NCBI GI
BLAST score
                   652
                   0.0e + 00
E value
                   986
Match length
                   93
% identity
                   G.max tufA gene for chloroplast translation elongation
NCBI Description
                   factor EF-Tu
                   1640
Seq. No.
                   893 4.R1040
Contig ID
                   LIB3170-014-Q1-K1-F12
 5'-most EST
Method
                   BLASTX
NCBI GI
                   g2500354
BLAST score
                   619
                   1.0e-111
E value
Match length
                   224
                   91
 % identity
                   60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj BAA19462
 NCBI Description
                   (AB001891) QM family protein [Solanum melongena]
```

Seq. No. 1641 Contig ID 893_5.R1040 5'-most EST hrw701059127.h1



```
BLASTX
Method
                  g121631
NCBI GI
                  308
BLAST score
                  3.0e-28
E value
Match length
                  74
                  77
% identity
                  GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 2 PRECURSOR
NCBI Description
                  >gi 72323_pir__KNNT2S glycine-rich protein 2 - wood tobacco
                  >gi_19743_emb_CAA42622_ (X60007) nsGRP-2 [Nicotiana
                  sylvestris]
                  1642
Seq. No.
                  893 6.R1040
Contig ID
                  fde700870636.hl
5'-most EST
Method
                  BLASTN
NCBI GI
                  g18775
                  218
BLAST score
                  1.0e-119
E value
                   420
Match length
                   97
% identity
                  G.max tufA gene for chloroplast translation elongation
NCBI Description
                   factor EF-Tu
                   1643
Seq. No.
                   893 7.R1040
Contig ID
                  LIB3039-016-Q1-E1-H9
5'-most EST
                   BLASTX
Method
                   g2500354
NCBI GI
                   550
BLAST score
                   2.0e-56
E value
                   109
Match length
                   94
% identity
                   60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894_dbj_BAA19462_
NCBI Description
                   (AB001891) QM family protein [Solanum melongena]
                   1644
Seq. No.
                   893 8.R1040
Contig ID
                   LIB3028-019-Q1-B1-F4
5'-most EST
                   1645
Seq. No.
                   893 9.R1040
Contig ID
                   LIB3106-013-Q1-K1-F4
5'-most EST
                   BLASTX
Method
                   g2500354
NCBI GI
BLAST score
                   213
E value
                   4.0e-17
                   46
Match length
                   85
% identity
                   60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj_BAA19462_
NCBI Description
                   (AB001891) QM family protein [Solanum melongena]
                   1646
Seq. No.
                   893 12.R1040
Contig ID
```

LIB3106-037-Q1-K1-D3 5'-most EST

Method BLASTX g2500354 NCBI GI 635 BLAST score



2.0e-66 E value 126 Match length 93 % identity 60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj_BAA19462_ NCBI Description (AB001891) QM family protein [Solanum melongena]

1647 Seq. No. 893 19.R1040 Contig ID dpv701103675.h1 5'-most EST Method BLASTX

g2500354 NCBI GI BLAST score 348 3.0e-33 E value Match length 63 % identity 98

60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj BAA19462 NCBI Description (AB001891) QM family protein [Solanum melongena]

1648 Seq. No. Contig ID

893 21.R1040 rlr700902438.hl 5'-most EST

BLASTX Method g2500354 NCBI GI 245 BLAST score 3.0e-21 E value Match length 48 90 % identity

60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj BAA19462 NCBI Description

(AB001891) QM family protein [Solanum melongena]

1649 Seq. No.

894 1.R1040 Contig ID

LIB3040-038-Q1-E1-G5 5'-most EST

BLASTX Method NCBI GI g3868758 844 BLAST score 1.0e-90 E value Match length 212 % identity 74

(D89802) elongation factor 1B gamma [Oryza sativa] NCBI Description

1650 Seq. No.

895 2.R1040 Contig ID

uC-gmropic019f08b1 5'-most EST

Seq. No. 1651

896 1.R1040 Contig ID k11701210028.h1 5'-most EST

Seq. No. 1652

898 2.R1040 Contig ID

LIB3039-021-Q1-E1-H5 5'-most EST

BLASTX Method g2191136 NCBI GI BLAST score 985 1.0e-134 E value 446 Match length



% identity (AF007269) Similar to UTP-Glucose Glucosyltransferase; NCBI Description

coded for by A. thaliana cDNA T46230; coded for by A. thaliana cDNA H76538; coded for by A. thaliana cDNA H76290

[Arabidopsis thaliana]

1653 Seq. No.

Contig ID 898 3.R1040 sat701012837.h1 5'-most EST

BLASTX Method g2191136 NCBI GI BLAST score 511 1.0e-87 E value 442 Match length % identity 32

(AF007269) Similar to UTP-Glucose Glucosyltransferase; NCBI Description

coded for by A. thaliana cDNA T46230; coded for by A.

thaliana cDNA H76538; coded for by A. thaliana cDNA H76290

[Arabidopsis thaliana]

Seq. No. 1654

898 4.R1040 Contig ID

LIB3040-018-Q1-E1-D4 5'-most EST

BLASTN Method g1638836 NCBI GI 324 BLAST score 0.0e + 00E value 1302 Match length 82 % identity

NCBI Description H.vulgare mRNA for alpha-tubulin

1655 Seq. No.

898 5.R1040 Contig ID

5'-most EST LIB3139-028-P1-N1-D1

Method BLASTN NCBI GI q20412 BLAST score 507 0.0e + 00E value 1307 Match length % identity

NCBI Description P.amygdalus mRNA for alpha-tubulin

Seq. No.

% identity

1656 898 6.R1040 Contig ID 5'-most EST g5606046 Method BLASTN g4098271 NCBI GI BLAST score 275 1.0e-152 E value Match length 1244

Triticum aestivum alpha-tubulin mRNA, complete cds NCBI Description

1657 Seq. No.

899 1.R1040 Contig ID 2DA-02-Q1-B1-B7 5'-most EST

83

384



```
1658
Seq. No.
                  900 3.R1040
Contig ID
                  uxk700672265.h1
5'-most EST
Method
                  BLASTN
                  g473216
NCBI GI
BLAST score
                  118
                  1.0e-59
E value
                  354
Match length
                  83
% identity
NCBI Description P.sativum (little marvel) HSC71.0 mRNA
                  1659
Seq. No.
                  900 4.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy170b02b1
Method
                  BLASTX
                  g1076746
NCBI GI
                  1030
BLAST score
                   0.0e + 00
E value
                  502
Match length
                   86
% identity
                  heat shock protein 70 - rice (fragment)
NCBI Description
                   >gi_763160_emb_CAA47948_ (X67711) heat shock protein 70
                   [Oryza sativa]
                   1660
Seq. No.
                   900 5.R1040
Contig ID
                   jC-gmst02400010c01a1
5'-most EST
                   BLASTX
Method
                   g629641
NCBI GI
BLAST score
                   1143
                   0.0e + 00
E value
                   440
Match length
% identity
                   PsHSC71.0 protein - garden pea >gi_1076530_pir__S53500
NCBI Description
                   HSC71.0 protein - garden pea >gi_473217_emb_CAA83548_
                   (Z32537) PsHSC71.0 [Pisum sativum]
                   1661
Seq. No.
                   900 6.R1040
Contia ID
                   wvk700683741.hl
5'-most EST
Method
                   BLASTX
NCBI GI
                   g629641
BLAST score
                   1158
E value
                   0.0e + 00
Match length
                   433
% identity
                   84
                   PsHSC71.0 protein - garden pea >gi 1076530_pir__S53500
NCBI Description
                   HSC71.0 protein - garden pea >gi 473217_emb_CAA83548_
                   (Z32537) PsHSC71.0 [Pisum sativum]
                   1662
Seq. No.
                   900 7.R1040
Contig ID
                   leu701146433.h1
5'-most EST
```

5'-most EST leu7011464

Method BLASTN

NCBI GI g473216

BLAST score 447

E value 0.0e+00



Match length 1860 % identity 87

NCBI Description P.sativum (little marvel) HSC71.0 mRNA

Seq. No. 1663

Contig ID 901 1.R1040

5'-most EST uC-gmronoir033b09b1

Method BLASTX
NCBI GI g2702277
BLAST score 397
E value 2.0e-38
Match length 109
% identity 71

NCBI Description (AC003033) putative cyclin g-associated kinase [Arabidopsis

thaliana] >gi 2914689 (AC003974) putative cyclin

g-associated kinase [Arabidopsis thaliana]

Seq. No. 1664

Contig ID 901_3.R1040 5'-most EST fua701040662.h1

Seq. No. 1665

Contig ID 902_1.R1040 5'-most EST 2DA-02-Q1-B1-C1

Seq. No. 1666

Contig ID 905_1.R1040 5'-most EST 2DA-02-Q1-B1-C12

Seq. No. 1667

Contig ID 906_1.R1040 5'-most EST 2DA-02-Q1-B1-C2

Seq. No. 1668

Contig ID 908 1.R1040

5'-most EST uC-gmrominsoy040a12b1

Method BLASTX
NCBI GI g2979555
BLAST score 340
E value 8.0e-32
Match length 150
% identity 49

NCBI Description (AC003680) unknown protein [Arabidopsis thaliana]

Seq. No. 1669

Contig ID 912_1.R1040 5'-most EST txt700736440.h1

Method BLASTX
NCBI GI g1350983
BLAST score 1082
E value 1.0e-118
Match length 234
% identity 87

NCBI Description 40S RIBOSOMAL PROTEIN S3A (CYC07 PROTEIN)

Seq. No. 1670

Contig ID 912_5.R1040



LIB3050-020-Q1-K1-C7 5'-most EST

BLASTX Method NCBI GI g1856971 BLAST score 272 E value 4.0e-24Match length 57 % identity

(D26058) This gene is specifically expressed at the S phase NCBI Description

during the cell cycle in the synchronous culture of

periwinkle cells. [Catharanthus roseus]

1671 Seq. No.

913 1.R1040 Contig ID seb700653717.h1 5'-most EST

BLASTX Method NCBI GI q3834307 BLAST score 739 E value 3.0e-78226 Match length % identity

(AC005679) Strong similarity to gene T10I14.120 gi_2832679 NCBI Description

putative protein from Arabidopsis thaliana BAC gb_AL021712.

ESTs gb_N65887 and gb_N65627 come from this gene.

[Arabidopsis thaliana]

1672 Seq. No.

913_2.R1040 Contig ID zzp700835039.h1 5'-most EST

Method BLASTX q3834307 NCBI GI 941 BLAST score 1.0e-102 E value 223 Match length 78 % identity

(AC005679) Strong similarity to gene T10I14.120 gi_2832679 NCBI Description

putative protein from Arabidopsis thaliana BAC gb AL021712.

ESTs gb_N65887 and gb_N65627 come from this gene.

[Arabidopsis thaliana]

1673 Seq. No.

913 3.R1040 Contig ID

LIB3106-099-Q1-K1-D3 5'-most EST

BLASTX Method q3834307 NCBI GI 326 BLAST score 3.0e-30 E value 79 Match length 78 % identity

(AC005679) Strong similarity to gene T10I14.120 gi_2832679 NCBI Description putative protein from Arabidopsis thaliana BAC gb AL021712.

ESTs gb N65887 and gb_N65627 come from this gene.

[Arabidopsis thaliana]

1674 Seq. No.

913 4.R1040 Contig ID 5'-most EST zsq701118614.hl

Method BLASTN



```
g3766106
NCBI GI
                  58
BLAST score
                  1.0e-23
E value
Match length
                  158
% identity
                  Arabidopsis thaliana chromosome 1 BAC F9K20 sequence,
NCBI Description
                  complete sequence [Arabidopsis thaliana]
                  1675
Seq. No.
                  913 5.R1040
Contig ID
                  leu701147327.h1
5'-most EST
                  BLASTX
Method
                  g3834307
NCBI GI
                   253
BLAST score
                   5.0e-22
E value
Match length
                   52
% identity
                   92
                   (AC005679) Strong similarity to gene T10I14.120 gi_2832679
NCBI Description
                   putative protein from Arabidopsis thaliana BAC gb AL021712.
                   ESTs gb N65887 and gb N65627 come from this gene.
                   [Arabidopsis thaliana]
                   1676
Seq. No.
                   914 1.R1040
Contig ID
                   dpv701102672.h1
5'-most EST
                   BLASTX
Method
                   g2213592
NCBI GI
                   822
BLAST score
                   9.0e-88
E value
                   287
Match length
                   37
% identity
                  (AC000348) T7N9.12 [Arabidopsis thaliana]
NCBI Description
                   1677
Seq. No.
                   914 2.R1040
Contig ID
                   sat701015380.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2213592
                   829
BLAST score
E value
                   1.0e-88
                   298
Match length
                   35
% identity
                   (AC000348) T7N9.12 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1678
                   914 3.R1040
Contig ID
5'-most EST
                   LIB3092-053-Q1-K1-B11
                   BLASTX
Method
NCBI GI
                   g2213592
BLAST score
                   291
                   3.0e-26
E value
Match length
                   85
 % identity
                   32
                   (AC000348) T7N9.12 [Arabidopsis thaliana]
```

1679 Seq. No.

NCBI Description

915 1.R1040 Contig ID



```
fde700871332.hl
5'-most EST
                  BLASTX
Method
                  q3600047
NCBI GI
                   476
BLAST score
                   1.0e-47
E value
                  160
Match length
% identity
                   (AF080120) similar to elongation factor EF-Ts [Arabidopsis
NCBI Description
                   thaliana]
                   1680
Seq. No.
                   917 1.R1040
Contig ID
                   zsg701119187.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1871577
                   732
BLAST score
                   2.0e-77
E value
                   168
Match length
                   82
% identity
                   (Y11553) putative 21kD protein precursor [Medicago sativa]
NCBI Description
                   1681
Seq. No.
                   917 2.R1040
Contig ID
                   uC-gmrominsoy256c09b1
5'-most EST
                   BLASTX
Method
                   g1871577
NCBI GI
                   656
BLAST score
                   1.0e-68
E value
                   194
Match length
                   68
% identity
                  (Y11553) putative 21kD protein precursor [Medicago sativa]
NCBI Description
                   1682
Seq. No.
                   917 3.R1040
Contig ID
                   LIB3139-102-P1-N1-G2
5'-most EST
                   BLASTX
Method
                   q1871577
NCBI GI
                   580
BLAST score
E value
                   8.0e-64
Match length
                   181
% identity
                   (Y11553) putative 21kD protein precursor [Medicago sativa]
NCBI Description
Seq. No.
                   1683
                   919 1.R1040
Contig ID
                   leu701151275.hl
5'-most EST
                   BLASTX
Method
                   g3914136
NCBI GI
BLAST score
                   306
                   1.0e-27
E value
Match length
                   116
                   54
% identity
                   NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (LTP)
```

1684 Seq. No.

NCBI Description

[Cicer arietinum]

>gi_2632171_emb_CAA05771_ (AJ002958) lipid transfer protein

NCBI GI

E value

BLAST score

Match length



```
919 3.R1040
Contig ID
                  LIB3040-061-Q1-E11-F5
5'-most EST
                  BLASTX
Method
                  q3914136
NCBI GI
                  127
BLAST score
                  1.0e-12
E value
                  92
Match length
                  43
% identity
                  NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (LTP)
NCBI Description
                  >gi 2632171_emb_CAA05771_ (AJ002958) lipid transfer protein
                  [Cicer arietinum]
                  1685
Seq. No.
Contig ID
                  920 1.R1040
                  sat701013038.h1
5'-most EST
                  BLASTX
Method
                  g548770
NCBI GI
                  1759
BLAST score
                  0.0e + 00
E value
Match length
                  385
                  83
% identity
                  60S RIBOSOMAL PROTEIN L3 >gi_481228_pir__S38359 ribosomal
NCBI Description
                  protein L3 - rice >gi_303853_dbj_BAA02155_ (D12630)
                  ribosomal protein L3 [Oryza sativa]
                   1686
Seq. No.
                   920 2.R1040
Contig ID
                   zsg701120409.h1
5'-most EST
                   BLASTX
Method
                   g132944
NCBI GI
                   296
BLAST score
E value
                   9.0e-27
Match length
                   61
                   89
% identity
                   60S RIBOSOMAL PROTEIN L3 >gi 81658 pir JQ0772 ribosomal
NCBI Description
                   protein L3 (ARP2) - Arabidopsis thaliana >gi_806279
                   (M32655) ribosomal protein [Arabidopsis thaliana]
Seq. No.
                   1687
                   920 3.R1040
Contig ID
                   seb700649013.hl
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2315210
BLAST score
                   113
                   2.0e-56
E value
Match length
                   325
% identity
                  Lycopersicon esculentum mRNA for proteasome, alpha subunit
NCBI Description
Seq. No.
                   1688
                   922 1.R1040
Contig ID
                   jC-\overline{g}mst02400058e10a1
5'-most EST
                   BLASTN
Method
```

g2924734 112

8.0e-56

371



 Seq. No.
 1689

 Contig ID
 922_2.R1040

 5'-most EST
 crh700852414.h1

 Method
 BLASTX

 NCBI GI
 g728880

NCBI GI g728880
BLAST score 183
E value 2.0e-13
Match length 55
% identity 62

NCBI Description N-TERMINAL ACETYLTRANSFERASE COMPLEX ARD1 SUBUNIT HOMOLOG >gi_517485_emb_CAA54691_ (X77588) ARD1 N-acetyl transferase homologue [Homo sapiens] >gi_1302661 (U52112) ARD1 N-acetyl

transferase related protein [Homo sapiens]

Seq. No. 1690

Contig ID 923_1.R1040 5'-most EST rlr700899522.h1

Method BLASTX
NCBI GI g2494113
BLAST score 778
E value 6.0e-83
Match length 192
% identity 75

NCBI Description (AC002376) Strong similarity to Musa pectate lyase

(gb X92943). ESTs gb_AA042458, gb_ATTS4502, gb_N38552 come

from this gene. [Arabidopsis thaliana]

Seq. No. 1691

Contig ID 924_2.R1040 5'-most EST kl1701207620.h1

Seq. No. 1692

Contig ID 925 1.R1040

5'-most EST LIB3051-106-Q1-K1-C8

Method BLASTX
NCBI GI g136057
BLAST score 980
E value 1.0e-106
Match length 253
% identity 74

NCBI Description TRIOSEPHOSPHATE ISOMERASE, CYTOSOLIC (TIM)

>gi_99499_pir__A32187 (S)-tetrahydroberberine oxidase Coptis japonica >gi 556171 (J04121) triosephosphate

isomerase [Coptis japonica]

Seq. No. 1693

Contig ID 925_2.R1040

5'-most EST LIB3109-003-Q1-K1-C9

Method BLASTX
NCBI GI g1351279
BLAST score 359
E value 2.0e-44
Match length 113



% identity TRIOSEPHOSPHATE ISOMERASE, CYTOSOLIC (TIM) NCBI Description >gi 602590 emb CAA58230 (X83227) triosephosphate isomerase [Petunia x hybrida] 1694 Seq. No. 925 3.R1040 Contig ID fC-gmle700743237f1 5'-most EST BLASTX Method g136063

NCBI GI BLAST score 181 E value 2.0e-13 Match length 68 % identity 48

TRIOSEPHOSPHATE ISOMERASE, CYTOSOLIC (TIM) NCBI Description

>gi 68426 pir__ISZMT triose-phosphate isomerase (EC 5.3.1.1) - maize >gi_168647 (L00371) triosephosphate isomerase 1 [Zea mays] >gi_217974_dbj_BAA00009_ (D00012)

triosephosphate isomerase [Zea mays]

1695 Seq. No. 925 4.R1040 Contig ID

5'-most EST LIB3051-108-Q1-K1-C11

Method BLASTX NCBI GI q136057 423 BLAST score 1.0e-41 E value Match length 112 % identity 72

TRIOSEPHOSPHATE ISOMERASE, CYTOSOLIC (TIM) NCBI Description

>gi_99499_pir__A32187 (S)-tetrahydroberberine oxidase -Coptis japonica >gi_556171 (J04121) triosephosphate

isomerase [Coptis japonica]

1696 Seq. No.

Contig ID 926 1.R1040 epx701105477.hl 5'-most EST

Method BLASTX NCBI GI g4127456 BLAST score 914 1.0e-98 E value Match length 232 % identity 78

(AJ010818) Cpn21 protein [Arabidopsis thaliana] NCBI Description

1697 Seq. No.

926 2.R1040 Contig ID

5'-most EST LIB3040-003-Q1-E1-D3

BLASTX Method q4127456 NCBI GI 343 BLAST score 3.0e-32 E value 96 Match length 49 % identity

(AJ010818) Cpn21 protein [Arabidopsis thaliana] NCBI Description

Seq. No. 1698



```
Contig ID
                   926 3.R1040
5'-most EST
                   LIB3138-052-Q1-N1-C12
Method
                   BLASTX
                   q3057150
NCBI GI
BLAST score
                   151
E value
                   6.0e-10
Match length
                   47
% identity
                   41
NCBI Description
                   (AF059037) chaperonin 10 [Arabidopsis thaliana]
Seq. No.
                   1699
Contig ID
                   928_1.R1040
                   2DA-02-Q1-B1-E4
5'-most EST
Method
                   BLASTN
NCBI GI
                   q2351069
BLAST score
                   57
E value
                   3.0e-23
Match length
                   227
% identity
                   81
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MSH12, complete sequence [Arabidopsis thaliana]
                   1700
Seq. No.
Contig ID
                   929 1.R1040
5'-most EST
                   gsv701055259.h1
Method
                   BLASTX
                   q2935523
NCBI GI
BLAST score
                   173
E value
                   4.0e-12
Match length
                   135
% identity
                   35
NCBI Description
                   (AF049066) 21 kD protein precursor [Pinus radiata]
Seq. No.
                   1701
Contig ID
                   929 2.R1040
5'-most EST
                   uC-gmropic059d02b1
Method
                   BLASTN
NCBI GI
                   g3821780
BLAST score
                   36
                   1.0e-10
E value
Match length
                   44
% identity
                   65
NCBI Description
                  Xenopus laevis cDNA clone 27A6-1
Seq. No.
                   1702
                   931 1.R1040
Contig ID
5'-most EST
                   vzy700751374.h1
Method
                  BLASTX
NCBI GI
                   g4559358
BLAST score
                   271
                   2.0e-23
E value
Match length
                   97
% identity
                   53
```

Seq. No. 1703

thaliana]

NCBI Description

(AC006585) putative steroid binding protein [Arabidopsis



```
931 2.R1040
Contig ID
5'-most EST
                  eep700868348.h1
                  BLASTX
Method
                  q4200044
NCBI GI
BLAST score
                  471
E value
                  4.0e-47
Match length
                  165
% identity
                  55
                  (AB022732) cytochrome P450 [Glycyrrhiza echinata]
NCBI Description
                  1704
Seq. No.
Contig ID
                  931 4.R1040
5'-most EST
                  rlr700900741.hl
Method
                  BLASTX
                  g4559358
NCBI GI
BLAST score
                  205
                  6.0e-16
E value
Match length
                  72
% identity
                  53
                  (AC006585) putative steroid binding protein [Arabidopsis
NCBI Description
                  thaliana]
                  1705
Seq. No.
Contig ID
                   932_1.R1040
                   jC-gmro02910036a06d1
5'-most EST
                  BLASTX
Method
                   q629602
NCBI GI
                   3123
BLAST score
E value
                   0.0e + 00
                   750
Match length
                   77
% identity
                  probable imbibition protein - wild cabbage
NCBI Description
                   >qi 488787 emb CAA55893 (X79330) putative imbibition
                   protein [Brassica oleracea]
Seq. No.
                   1706
                   932 2.R1040
Contig ID
                   q5606124
5'-most EST
Method
                   BLASTX
                   g629602
NCBI GI
                   587
BLAST score
                   3.0e-60
E value
                   142
Match length
                   76
% identity
                  probable imbibition protein - wild cabbage
NCBI Description
                   >gi 488787 emb CAA55893 (X79330) putative imbibition
                   protein [Brassica oleracea]
Seq. No.
                   1707
                   932 3.R1040
Contig ID
5'-most EST
                   LIB3170-050-Q1-K1-D8
Seq. No.
                   1708
```

Contig ID 933_1.R1040 5'-most EST 2DA-02-Q1-B1-E9

Method BLASTX NCBI GI g3367569



```
BLAST score
                  175
                  4.0e-12
E value
Match length
                  173
% identity
                  31
                  (AL031135) putative protein [Arabidopsis thaliana]
NCBI Description
                  1709
Seq. No.
                  934 1.R1040
Contig ID
5'-most EST
                  leu701144440.h1
Method
                  BLASTX
                  g2832617
NCBI GI
BLAST score
                  256
                  4.0e-22
E value
Match length
                  71
% identity
                  66
                  (AL021711) putative protein [Arabidopsis thaliana]
NCBI Description
                  1710
Seq. No.
                  936 1.R1040
Contig ID
                  hrw701057246.h1
5'-most EST
                  1711
Seq. No.
                  938 1.R1040
Contig ID
                  2DA-02-Q1-B1-F2
5'-most EST
Method
                  BLASTX
                  q3077640
NCBI GI
BLAST score
                  304
                  9.0e-28
E value
Match length
                  116
% identity
                  55
                  (AJ223151) O-methyltransferase [Prunus dulcis]
NCBI Description
Seq. No.
                  1712
Contig ID
                   941 1.R1040
5'-most EST
                  LIB3170-085-Q1-K1-G6
Method
                  BLASTX
NCBI GI
                  g2447107
BLAST score
                   328
E value
                   3.0e-30
Match length
                  172
% identity
                   42
NCBI Description
                  (U42580) A638R [Paramecium bursaria Chlorella virus 1]
Seq. No.
                  1713
Contig ID
                   942 1.R1040
5'-most EST
                   2DA-02-Q1-B1-F6
Seq. No.
                  1714
                   942 2.R1040
Contig ID
5'-most EST
                  rlr700899163.h1
```

Seq. No. 1715

Contig ID 943_1.R1040

5'-most EST LIB3039-005-Q1-E1-G3

Method BLASTX NCBI GI g125722 BLAST score 365



E value 1.0e-34
Match length 228
% identity 44

NCBI Description KUNITZ-TYPE TRYPSIN INHIBITOR KTI1 PRECURSOR

>gi_81814_pir__JQ1091 trypsin inhibitor KTi1 (Kunitz) soybean >gi_256635_bbs_115028 (S45035) Kunitz trypsin
inhibitor KTi1 [soybeans, Peptide, 203 aa] [Glycine max]

Seq. No. 1716

Contig ID 944 1.R1040 5'-most EST pmv700888263.h1

Method BLASTX
NCBI GI g3522939
BLAST score 411
E value 6.0e-55
Match length 252
% identity 52

NCBI Description (AC004411) putative squamosa-promoter binding protein

[Arabidopsis thaliana]

Seq. No. 1717

Contig ID 945 1.R1040

5'-most EST LIB3093-036-Q1-K1-F1

Method BLASTX
NCBI GI g2982297
BLAST score 1108
E value 1.0e-121
Match length 232
% identity 93

NCBI Description (AF051233) KIAA0107-like protein [Picea mariana]

Seq. No. 1718

Contig ID 945_2.R1040 5'-most EST wrg700787648.h2

Method BLASTX
NCBI GI g2494625
BLAST score 187
E value 5.0e-14
Match length 109
% identity 40

NCBI Description HYPOTHETICAL PROTEIN KIAA0107 >qi 285951 dbj BAA03497

(D14663) KIAA0107 [Homo sapiens]

Seq. No. 1719

Contig ID 947 2.R1040

5'-most EST jC-gmro02910059e07a1

Method BLASTX
NCBI GI g3851636
BLAST score 471
E value 3.0e-47
Match length 108
% identity 83

NCBI Description (AF098519) unknown [Avicennia marina] >gi_4128206

(AF056316) 40S ribosome protein S7 [Avicennia marina]

Seq. No. 1720

Contig ID 947_3.R1040

```
5'-most EST
                  q5126873
                  BLASTX
Method
NCBI GI
                  g4426565
BLAST score
                  145
                  7.0e-09
E value
                  66
Match length
                  39
% identity
NCBI Description
                  (AF031483) unknown [Rattus norvegicus]
                  1721
Seq. No.
                  947 4.R1040
Contig ID
5'-most EST
                  LIB3093-014-Q1-K1-B9
Method
                  BLASTX
                  g4426565
NCBI GI
BLAST score
                  175
E value
                  8.0e-12
Match length
                  238
                  34
% identity
NCBI Description (AF031483) unknown [Rattus norvegicus]
                  1722
Seq. No.
                  947 5.R1040
Contig ID
5'-most EST
                  LIB3074-003-Q1-K1-B7
Method
                  BLASTX
NCBI GI
                  q126085
BLAST score
                  242
                  5.0e-26
E value
                  236
Match length
% identity
                  28
                  SEED LECTIN SUBUNITS I AND II PRECURSOR
NCBI Description
                  >gi 81750_pir__A29572 lectin precursor - horse gram
                  >gi_167564 (J02721) lectin subunit I precursor [Dolichos
                  biflorus]
Seq. No.
                  1723
Contig ID
                  947 6.R1040
                  ncj700986678.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  q3851636
BLAST score
                  255
                  9.0e-22
E value
                  58
Match length
% identity
                  83
NCBI Description
                  (AF098519) unknown [Avicennia marina] >gi 4128206
                   (AF056316) 40S ribosome protein S7 [Avicennia marina]
                  1724
Seq. No.
Contig ID
                  948 1.R1040
```

5'-most EST jC-gmst02400054d09a1

Seq. No.

1725

948 2.R1040 Contig ID 5'-most EST asn701132457.h1

Seq. No.

1726

Contig ID

950 1.R1040

5'-most EST LIB3050-005-Q1-K1-C10



```
BLASTX
Method
NCBI GI
                  q4325282
                   584
BLAST score
                  7.0e-60
E value
                   158
Match length
                   70
% identity
                   (AF123310) NAC domain protein NAM [Arabidopsis thaliana]
NCBI Description
                   >gi_4325286_gb_AAD17314_ (AF123311) NAC domain protein NAM
                   [Arabidopsis thaliana]
                   1727
Seq. No.
Contig ID
                   950 2.R1040
5'-most EST
                   jC-qmle01810020d08a1
                   BLASTX
Method
                   a4218535
NCBI GI
                   591
BLAST score
                   7.0e-61
E value
Match length
                   167
% identity
                   63
                  (AJ010829) GRAB1 protein [Triticum sp.]
NCBI Description
                   1728
Seq. No.
Contig ID
                   950 3.R1040
                   zzp700834856.h1
5'-most EST
                   BLASTX
Method
                   q2058282
NCBI GI
BLAST score
                   716
E value
                   2.0e-75
Match length
                   177
                   78
% identity
NCBI Description (X97377) atranbpla [Arabidopsis thaliana]
Seq. No.
                   1729
                   950 4.R1040
Contig ID
                   leu701152450.h1
5'-most EST
                   1730
Seq. No.
                   950 5.R1040
Contig ID
                   LIB3030-006-Q1-B1-A1
5'-most EST
                   BLASTX
Method
                   g4218535
NCBI GI
                   525
BLAST score
                   3.0e-53
E value
                   144
Match length
                   65
% identity
                   (AJ010829) GRAB1 protein [Triticum sp.]
NCBI Description
Seq. No.
                   1731
                   950 6.R1040
Contig ID
5'-most EST
                   pcp700992258.h1
                   BLASTX
Method
NCBI GI
                   g2443890
BLAST score
                   221
                   1.0e-17
E value
Match length
                   66
```

NCBI Description (AC002294) similar to NAM (gp_X92205_1321924) and CUC2

68

% identity



(gp AB002560 1944132) proteins [Arabidopsis thaliana]

Seq. No. 1732

Contig ID 950 9.R1040

5'-most EST LIB3092-043-Q1-K1-E1

Method BLASTX
NCBI GI 94218535
BLAST score 374
E value 6.0e-36
Match length 93
% identity 70

NCBI Description (AJ010829) GRAB1 protein [Triticum sp.]

Seq. No. 1733

Contig ID 950 10.R1040

5'-most EST jC-gmst02400071g08d1

Seq. No. 1734

Contig ID 950 11.R1040

5'-most EST uC-gmflminsoy049c06b1

Seq. No. 1735

Contig ID 950 12.R1040

5'-most EST LIB3074-028-Q1-K1-C3

Seq. No. 1736

Contig ID 950_14.R1040 5'-most EST zsg701127580.h1

Seq. No. 1737

Contig ID 950 15.R1040

5'-most EST LIB3039-012-Q1-E1-H3

Method BLASTX
NCBI GI g1732511
BLAST score 511
E value 6.0e-52
Match length 121
% identity 79

NCBI Description (U62742) Ran binding protein 1 homolog [Arabidopsis

thaliana]

Seq. No. 1738

Contig ID 950 16.R1040

5'-most EST LIB3073-009-Q1-K1-C4

Seq. No. 1739

Contig ID 950_19.R1040

5'-most EST LIB3106-112-Q1-K1-A7

Method BLASTN
NCBI GI g1336081
BLAST score 45
E value 4.0e-16
Match length 77
% identity 90

NCBI Description Glycine max var. Century ascorbate peroxidase 2 (APx2)

mRNA, complete cds



Seq. No. 1740

Contig ID 950_24.R1040

5'-most EST LIB3106-061-Q1-K1-G2

Seq. No. 1741

Contig ID 950_25.R1040 5'-most EST hyd700726095.h1

Seq. No. 1742

Contig ID 951 1.R1040

5'-most EST LIB3039-026-Q1-E1-H3

Method BLASTX
NCBI GI g1173209
BLAST score 681
E value 1.0e-71
Match length 140
% identity 94

NCBI Description 40S RIBOSOMAL PROTEIN S16 >gi_541835_pir__S41193 ribosomal

protein S16 protein - upland cotton

>gi_439654_emb_CAA53567_ (X75954) RS16 protein, 40S subunit

[Gossypium hirsutum]

Seq. No. 1743

Contig ID 951_2.R1040 5'-most EST dpv701100901.h1

Method BLASTX
NCBI GI g1173209
BLAST score 683
E value 8.0e-72
Match length 140
% identity 95

NCBI Description 40S RIBOSOMAL PROTEIN S16 >gi_541835_pir__S41193 ribosomal

protein S16 protein - upland cotton

>gi_439654_emb_CAA53567_ (X75954) RS16 protein, 40S subunit

[Gossypium hirsutum]

Seq. No. 1744

Contig ID 951_3.R1040

5'-most EST jC-gmst02400009b06a1

Method BLASTX
NCBI GI g1173209
BLAST score 592
E value 2.0e-61
Match length 125
% identity 93

NCBI Description 40S RIBOSOMAL PROTEIN S16 >gi_541835_pir__S41193 ribosomal

protein S16 protein - upland cotton

>gi_439654_emb_CAA53567_ (X75954) RS16 protein, 40S subunit

·** - -

[Gossypium hirsutum]

Seq. No. 1745

Contig ID 954_1.R1040 5'-most EST sat701007156.h1

Method BLASTX
NCBI GI g1486472
BLAST score 963
E value 1.0e-107



Match length 240 % identity 84

NCBI Description (X99853) oxoglutarate malate translocator [Solanum

tuberosum]

Seq. No. 1746

Contig ID 954_2.R1040 5'-most EST gsv701045595.h1

Seq. No. 1747

Contig ID 954 3.R1040

5'-most EST LIB3107-070-Q1-K1-G12

Method BLASTX
NCBI GI g1486472
BLAST score 625
E value 4.0e-65
Match length 135
% identity 90

NCBI Description (X99853) oxoglutarate malate translocator [Solanum

tuberosum]

Seq. No. 1748

Contig ID 954 4.R1040

5'-most EST uC-gmropic109e10b1

Method BLASTX
NCBI GI g2244847
BLAST score 1400
E value 1.0e-155
Match length 529
% identity 56

NCBI Description (Z97337) hydroxyproline-rich glycoprotein homolog

[Arabidopsis thaliana]

Seq. No. 1749

Contig ID 954 5.R1040

5'-most EST LIB3109-052-Q1-K1-D9

Seq. No. 1750

Contig ID 954_6.R1040

5'-most EST jC-gmle01810016e03a1

Seq. No. 1751

Contig ID 955 1.R1040

5'-most EST LIB3065-024-Q1-N1-B5

Seq. No. 1752

Contig ID 956_1.R1040

5'-most EST LIB3138-047-Q1-N1-D3

Method BLASTX
NCBI GI g3513736
BLAST score 311
E value 2.0e-28
Match length 141
% identity 49

NCBI Description (AF080118) No definition line found [Arabidopsis thaliana]

>gi_4539367_emb_CAB40061.1_ (AL049525) putative protein

[Arabidopsis thaliana]

```
1753
Seq. No.
                   957 1.R1040
Contig ID
5'-most EST
                   LIB3109-028-Q1-K1-G2
Method
                   BLASTX
                   q606942
NCBI GI
BLAST score
                   577
                   3.0e-59
E value
                   190
Match length
% identity
                   60
```

(U13760) unknown [Gossypium hirsutum] NCBI Description

1754 Seq. No.

958 1.R1040 Contig ID

5'-most EST LIB3093-033-Q1-K1-C6

Method BLASTX g120658 NCBI GI BLAST score 1645 0.0e+00E value 403 Match length % identity 82

GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE A PRECURSOR, NCBI Description

CHLOROPLAST >gi 66025_pir__DEPMNA

glyceraldehyde-3-phosphate dehydrogenase (NADP+)

(phosphorylating) (EC 1.2.1.13) A precursor, chloroplast garden pea >gi 12159 emb_CAA36396_ (X52148) GAPA [Pisum

sativum]

Seq. No. 1755

958 2.R1040 Contig ID taw700658860.h1 5'-most EST

BLASTN Method g20728 NCBI GI BLAST score 37 1.0e-11 E value 68 Match length 91 % identity

Pea chloroplast GAPA mRNA encoding NCBI Description

glyceraldehyde-3-phosphate dehydrogenase (GAPDH) subunit A

(EC 1.2.1.13)

1756 Seq. No.

960 1.R1040 Contig ID

LIB3170-079-Q1-K1-A4 5'-most EST

Method BLASTX q585749 NCBI GI BLAST score 718 8.0e-76 E value Match length 177 79 % identity

PHYTOENE SYNTHASE PRECURSOR >gi 433994 emb CAA48155 NCBI Description

(X68017) phytoene synthase [Capsicum annuum]

1757 Seq. No.

Contig ID 960 2.R1040 vzy700755950.h1 5'-most EST

BLASTX Method

```
g585749
NCBI GI
                  259
BLAST score
                  5.0e-23
E value
Match length
                  105
% identity
                  61
                  PHYTOENE SYNTHASE PRECURSOR >gi 433994 emb CAA48155_
NCBI Description
                   (X68017) phytoene synthase [Capsicum annuum]
                  1758
Seq. No.
                  961 1.R1040
Contig ID
                  LIB3170-087-Q1-K1-C2
5'-most EST
                  1759
Seq. No.
                  961 2.R1040
Contig ID
5'-most EST
                  LIB3170-084-Q1-J1-B1
                  1760
Seq. No.
                  962 1.R1040
Contig ID
                  LIB3109-010-Q1-K1-E9
5'-most EST
                  BLASTX
Method
                  g1871199
NCBI GI
BLAST score
                  189
                   4.0e-14
E value
                  103
Match length
                   40
% identity
                  (U91318) pM5 (3' partial) [Homo sapiens]
NCBI Description
                   1761
Seq. No.
                   963 1.R1040
Contig ID
5'-most EST
                  LIB3040-004-Q1-E1-C8
                  BLASTX
Method
                   g2501356
NCBI GI
BLAST score
                   3168
E value
                   0.0e+00
Match length
                   738
                   83
% identity
                   TRANSKETOLASE, CHLOROPLAST PRECURSOR (TK)
NCBI Description
                   >gi 1658322 emb CAA90427_ (Z50099) transketolase precursor
                   [Solanum tuberosum]
                   1762
Seq. No.
                   963 2.R1040
Contig ID
                   q5606060
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2501353
BLAST score
                   608
                   5.0e-63
E value
Match length
                   140
% identity
                   TRANSKETOLASE, CHLOROPLAST (TK) >gi_1084440_pir__S54300
NCBI Description
```

transketolase (EC 2.2.1.1) 3 - Craterostigma plantagineum (fragment) >gi 664901 emb CAA86607 (Z46646) transketolase

[Craterostigma plantagineum]

Seq. No.

1763

963 3.R1040 Contig ID

LIB3109-035-Q1-K1-G8 5'-most EST



Method BLASTX
NCBI GI g2501356
BLAST score 165
E value 3.0e-11
Match length 97
% identity 37

NCBI Description TRANSKETOLASE, CHLOROPLAST PRECURSOR (TK)

>gi_1658322_emb_CAA90427_ (Z50099) transketolase precursor

[Solanum tuberosum]

Seq. No. 1764

Contig ID 964 1.R1040 5'-most EST sat701014744.h1

Method BLASTX
NCBI GI g3341688
BLAST score 1023
E value 1.0e-111
Match length 259
% identity 74

NCBI Description (AC003672) putative casein kinase II beta subunit

[Arabidopsis thaliana]

Seq. No. 1765

Contig ID 964_2.R1040

5'-most EST LIB3093-018-Q1-K1-A6

Method BLASTX
NCBI GI g729880
BLAST score 235
E value 2.0e-19
Match length 124
% identity 46

NCBI Description CASEIN KINASE II BETA CHAIN (CK II) >gi_1076299_pir_S47967

casein kinase II (EC 2.7.1.-) beta chain CKB1 - Arabidopsis thaliana >gi_468264 (L22563) casein kinase II beta subunit

CKB1 [Arabidopsis thaliana]

Seq. No. 1766

Contig ID 964 3.R1040

5'-most EST LIB3092-047-Q1-K1-B2

Method BLASTX
NCBI GI 9729880
BLAST score 168
E value 1.0e-11
Match length 113
% identity 41

NCBI Description CASEIN KINASE II BETA CHAIN (CK II) >gi_1076299_pir__S47967

casein kinase II (EC 2.7.1.-) beta chain CKB1 - Arabidopsis thaliana >gi 468264 (L22563) casein kinase II beta subunit

CKB1 [Arabidopsis thaliana]

Seq. No. 1767

Contig ID 964_4.R1040 5'-most EST jex700908130.h1

Method BLASTN
NCBI GI g468263
BLAST score 50
E value 5.0e-19



122 Match length 85 % identity

Arabidopsis thaliana casein kinase II beta subunit CKB1 NCBI Description

mRNA, complete cds

Seq. No. Contig ID

5'-most EST

1768 967 1.R1040 q5677191

Seq. No.

1769 968 1.R1040

Contig ID 5'-most EST

 $jC-\overline{g}mst02400015g09d1$

Seq. No.

1770

Contig ID

970 1.R1040

5'-most EST

 $jC-\overline{g}mle01810044a05d1$

Method NCBI GI BLAST score BLASTX q4006868 382

E value Match length 9.0e-37 110 65

% identity NCBI Description

(Z99707) putative protein [Arabidopsis thaliana]

Seq. No. Contig ID 1771

974 1.R1040 hyd700727370.h1 5'-most EST

Seq. No.

1772

Contig ID

975 1.R1040

5'-most EST

jC-gmro02910008f01a1

Method NCBI GI BLAST score BLASTX g687844 982

E value Match length 1.0e-106

% identity

629 19

NCBI Description

(U21320) contains TPR domain-like repeats [Caenorhabditis

elegans]

Seq. No.

1773

Contig ID

975 2.R1040

5'-most EST

LIB3106-034-Q1-K1-D4

Seq. No.

1774

Contig ID 5'-most EST

975 3.R1040 pmv700888650.h1

Seq. No. Contig ID 1775

975 4.R1040

5'-most EST

LIB3170-043-Q1-K1-A4

Seq. No. Contig ID 1776

5'-most EST

975 5.R1040 LIB3106-074-Q1-K1-G9

Method NCBI GI BLASTX

g2266994



BLAST score 6.0e-16 E value 133 Match length % identity NCBI Description

(U77413) O-linked GlcNAc transferase [Homo sapiens] >gi_4505499_ref_NP_003596.1_pOGT_ O-GlcNAc transferase (uridine diphospho-N-acetylglucosamine:polypeptide

beta-N-acetylglucosaminyl transferase)

1777 Seq. No.

977 1.R1040 Contig ID

ASG32442DA-01-Q1-E1-A8 5'-most EST

BLASTX Method q4538945 NCBI GI BLAST score 361 E value 3.0e-34 Match length 82 % identity 78

(AL049483) putative thioredoxin [Arabidopsis thaliana] NCBI Description

1778 Seq. No.

984 1.R1040 Contig ID

5'-most EST LIB3170-052-Q1-J1-E12

1779 Seq. No.

984_3.R1040 Contig ID

 $uC\text{--}\overline{g}\text{mrominsoy}133\text{h}03\text{b}1$ 5'-most EST

Seq. No. 1780

985 1.R1040 Contig ID

LIB3138-123-Q1-N1-D12 5'-most EST

1781 Seq. No.

986 1.R1040 Contig ID g5606936 5'-most EST

1782 Seq. No.

988 1.R1040 Contig ID

LIB3051-046-Q1-K1-B6 5'-most EST

1783 Seq. No.

989 1.R1040 Contig ID

ASG32442DA-01-Q1-E1-C1 5'-most EST

1784 Seq. No.

992 1.R1040 Contig ID g5509751 5'-most EST

1785 Seq. No.

995 1.R1040 Contig ID

jC-qmst02400016d12d1 5'-most EST

BLASTN Method q3168616 NCBI GI BLAST score 75 5.0e-34E value Match length 119 % identity 91



NCBI Description Drosophila melanogaster DNA sequence (P1 DS07982 (D185)), complete sequence [Drosophila melanogaster]

Seq. No. 1786

Contig ID 997_1.R1040

5'-most EST ASG32442DA-01-Q1-E1-C7

Seq. No. 1787

Contig ID 1004 1.R1040

5'-most EST LIB3028-017-Q1-B1-F12

Seq. No. 1788

Contig ID 1005_1.R1040

5'-most EST LIB3053-012-Q1-N1-A6

Seq. No. 1789

Contig ID 1010 1.R1040

5'-most EST jC-gmle01810088b06d1

Method BLASTN
NCBI GI g12139
BLAST score 387
E value 0.0e+00
Match length 499
% identity 94

NCBI Description Pea plastid genes rps2, atpI, atpH, atpF, atpA, trnR and trnG coding for ribosomal protein S2, one CF(1) and three

CF(O) subunits of ATP synthase and tRNA-Arg and tRNA-Gly

Seq. No. 1790

Contig ID 1015_1.R1040

5'-most EST ASG32442DA-01-Q1-E1-E6

Seq. No. 1791

Contig ID 1019_1.R1040

5'-most EST jC-gmro02910031e05d1

Method BLASTN
NCBI GI g2598578
BLAST score 95
E value 5.0e-46
Match length 171
% identity 89

NCBI Description Medicago truncatula mRNA for L-ascorbate oxidase

Seq. No. 1792

Contig ID 1025_1.R1040

5'-most EST ASG32442DA-01-Q1-E1-F4

Seq. No. 1793

Contig ID 1032_1.R1040

5'-most EST LIB3170-087-Q1-J1-B2

Method BLASTX
NCBI GI 94512703
BLAST score 442
E value 2.0e-43
Match length 211
% identity 45

NCBI Description (AC006569) unknown protein [Arabidopsis thaliana]



```
1794
Seq. No.
                   1035 1.R1040
Contig ID
                   ASG32442DA-01-Q1-E1-G4
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2129854
                   418
BLAST score
                   4.0e-41
E value
                   105
Match length
                   72
% identity
                   early nodulin 8 precursor - alfalfa >gi_304037 (L18899)
NCBI Description
                   early nodulin [Medicago sativa]
                   1795
Seq. No.
                   1046 1.R1040
Contig ID
5'-most EST
                   q4299082
                   BLASTX
Method
NCBI GI
                   q4510383
                   762
BLAST score
                   9.0e-81
E value
                   303
Match length
% identity
                   (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
                   1796
Seq. No.
                   1046 2.R1040
Contig ID
5'-most EST
                   q430\overline{5}468
                   BLASTX
Method
                   q4510383
NCBI GI
BLAST score
                   271
                   1.0e-23
E value
                   82
Match length
% identity
                    67
                   (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    1797
                    1050 1.R1040
Contig ID
                    \verb"jC-gmro02910007g06d1"
5'-most EST
                    BLASTX
Method
                    g3953471
NCBI GI
                    1320
BLAST score
                    1.0e-146
E value
                    311
Match length
                    79
 % identity
                   (AC002328) F2202.16 [Arabidopsis thaliana]
 NCBI Description
                    1798
 Seq. No.
                    1050 2.R1040
 Contig ID
                    LIB3170-029-Q1-J1-A7
 5'-most EST
                    BLASTX
 Method
 NCBI GI
                    a3953471
 BLAST score
                    249
                    3.0e-21
 E value
                    58
 Match length
```

(AC002328) F2202.16 [Arabidopsis thaliana]

76

% identity

NCBI Description

BLAST score

% identity

E value Match length 1566 1.0e-175

414

71



```
1799
Seq. No.
                  1050 3.R1040
Contig ID
                  LIB3030-003-Q1-B1-B12
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3953471
BLAST score
                  624
                  7.0e-65
E value
Match length
                  231
% identity
NCBI Description (AC002328) F2202.16 [Arabidopsis thaliana]
                  1800
Seq. No.
Contig ID
                  1050 4.R1040
5'-most EST
                  ASG32442DA-02-Q1-E1-A11
Method
                  BLASTX
NCBI GI
                  g3953471
BLAST score
                  229
E value
                  6.0e-19
                  50
Match length
% identity
                  80
NCBI Description (AC002328) F2202.16 [Arabidopsis thaliana]
                  1801
Seq. No.
Contig ID
                  1051 1.R1040
                  g4285505
5'-most EST
                  1802
Seq. No.
Contig ID
                  1057 1.R1040
                  ASG32442DA-02-Q1-E1-A7
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1652704
BLAST score
                  198
                  4.0e-15
E value
                  95
Match length
                  41
% identity
NCBI Description (D90907) hypothetical protein [Synechocystis sp.]
Seq. No.
                  1803
Contig ID
                  1060 1.R1040
5'-most EST
                  ASG32442DA-02-Q1-E1-B1
Method
                  BLASTX
NCBI GI
                  g1076288
BLAST score
                  241
                  3.0e-20
E value
Match length
                  89
                  48
% identity
NCBI Description amino acid permease AAP3 - Arabidopsis thaliana
Seq. No.
                  1804
Contig ID
                  1064_1.R1040
5'-most EST
                  jC-gmle01810031g05d1
Method
                  BLASTX
NCBI GI
                  q3868758
```



```
(D89802) elongation factor 1B gamma [Oryza sativa]
NCBI Description
                  1805
Seq. No.
Contig ID
                  1064 2.R1040
5'-most EST
                  g5126554
                  BLASTX
Method
                   q3868758
NCBI GI
                   518
BLAST score
                   1.0e-52
E value
Match length
                   144
% identity
                   68
                   (D89802) elongation factor 1B gamma [Oryza sativa]
NCBI Description
                   1806
Seq. No.
Contig ID
                   1064 4.R1040
                   LIB3139-055-P1-N1-C4
5'-most EST
                   BLASTX
Method
                   q3868758
NCBI GI
                   420
BLAST score
                   2.0e-41
E value
                   104
Match length
                   75
% identity
                   (D89802) elongation factor 1B gamma [Oryza sativa]
NCBI Description
                   1807
Seq. No.
                   1064 5.R1040
Contig ID
                   LIB3170-036-Q1-K2-D11
5'-most EST
                   BLASTX
Method
                   g2160158
NCBI GI
                   212
BLAST score
                   4.0e-17
E value
                   43
Match length
                   93
% identity
                   (AC000132) Similar to elongation factor 1-gamma
NCBI Description
                   (gb EF1G_XENLA). ESTs gb_T20564,gb_T45940,gb_T04527 come
                   from this gene. [Arabidopsis thaliana]
                   1808
Seq. No.
                   1064 6.R1040
Contig ID
5'-most EST
                   sat701004617.h1
                   BLASTN
Method
                   q3868757
NCBI GI
BLAST score
                   73
E value
                   5.0e-33
Match length
                   209
                   84
% identity
                   Oryza sativa mRNA for elongation factor 1B gamma, complete
NCBI Description
                   1809
 Seq. No.
                   1065 1.R1040
 Contig ID
                   ASG32442DA-02-Q1-E1-B3
 5'-most EST
                   BLASTX
 Method
                   g4454482
 NCBI GI
                    410
 BLAST score
```

5.0e-40

96

E value

Match length



```
% identity
NCBI Description (AC006234) hypothetical protein [Arabidopsis thaliana]
                   1810
Seq. No.
                   1071 1.R1040
Contig ID
                   LIB3093-026-Q1-K1-H10
5'-most EST
                   1811
Seq. No.
                   1074 1.R1040
Contig ID
                   jC-g\overline{m}le01810087f04d1
5'-most EST
                   1812
Seq. No.
                   1074 2.R1040
Contig ID
                   jC-gmle01810046b09d1
5'-most EST
Seq. No.
                   1813
```

1092 1.R1040 Contig ID jC-gmst02400008f02a1 5'-most EST

1814 Seq. No.

1092 2.R1040 Contig ID

ASG32442DA-02-Q1-E1-D8 5'-most EST

1815 Seq. No.

1093 1.R1040 Contig ID

LIB3051-013-Q1-E1-A4 5'-most EST

Seq. No. 1816

1094 1.R1040 Contig ID

ASG32442DA-02-Q1-E1-E1 5'-most EST

BLASTX Method g3063453 NCBI GI 167 BLAST score 2.0e-11 E value 147 Match length 36 % identity

NCBI Description (AC003981) F22013.15 [Arabidopsis thaliana]

1817 Seq. No.

1098 1.R1040 Contig ID

ASG32442DA-02-Q1-E1-E3 5'-most EST

BLASTX Method q4567210 NCBI GI 548 BLAST score 1.0e-55 E value Match length 110 93 % identity

(AC007168) unknown protein [Arabidopsis thaliana] NCBI Description

1818 Seq. No.

1098 2.R1040 Contig ID

jC-qmst02400058b05d1 5'-most EST

BLASTX Method g4567210 NCBI GI BLAST score 191 3.0e-14E value Match length 61



```
% identity
                   (AC007168) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1819
                   1100 1.R1040
Contig ID
                   jC-gmf102220114f08d1
5'-most EST
                   BLASTX
Method
                   q4176531
NCBI GI
                   150
BLAST score
                   2.0e-09
E value
                   92
Match length
                   37
% identity
                   (AL035263) weak similarity to chick phosphatidylcholine-ste
NCBI Description
                   rol acetyltransferase [Schizosaccharomyces pombe]
                   1820
Seq. No.
                   1101 1.R1040
Contig ID
                   ASG3\overline{2}442DA-02-Q1-E1-E5
5'-most EST
                   BLASTX
Method
                   g2098709
NCBI GI
                   473
BLAST score
                   8.0e-59
E value
                   191
Match length
                   58
% identity
                  (U82975) pectinesterase [Citrus sinensis]
NCBI Description
                   1821
Seq. No.
                   1103 1.R1040
Contig ID
                   LIB3093-001-Q1-K1-G8
5'-most EST
                   BLASTN
Method
                   g3821780
NCBI GI
                   36
BLAST score
                   1.0e-10
E value
                   60
Match length
                   56
% identity
NCBI Description Xenopus laevis cDNA clone 27A6-1
                   1822
Seq. No.
                   1103 3.R1040
Contig ID
5'-most EST
                   jC-gmf102220133h03d1
Seq. No.
                   1823
                   1112 1.R1040
Contig ID
                   jC-gmst02400009a05d1
5'-most EST
Method
                   BLASTX
                   g2447107
NCBI GI
BLAST score
                   302
E value
                   4.0e-27
                   125
Match length
                   50
 % identity
                   (U42580) A638R [Paramecium bursaria Chlorella virus 1]
NCBI Description
                   1824
 Seq. No.
                   1114 1.R1040
 Contig ID
                   g5509228
 5'-most EST
```

BLASTX

g4160280

Method

NCBI GI



```
1489
BLAST score
                  1.0e-166
E value
                  430
Match length
                  62
% identity
                  (AJ006224) purple acid phosphatase [Ipomoea batatas]
NCBI Description
                  1825
Seq. No.
                  1114 2.R1040
Contig ID
                  jC-gmle01810014f06d1
5'-most EST
                  BLASTX
Method
                  g4160280
NCBI GI
                   290
BLAST score
                   6.0e-26
E value
                   76
Match length
% identity
                   (AJ006224) purple acid phosphatase [Ipomoea batatas]
NCBI Description
                   1826
Seq. No.
                   1114 5.R1040
Contig ID
                   LIB3106-010-Q1-K1-F9
5'-most EST
                   BLASTX
Method
                   g4160280
NCBI GI
                   280
BLAST score
                   6.0e-25
E value
                   90
Match length
                   59
% identity
                  (AJ006224) purple acid phosphatase [Ipomoea batatas]
NCBI Description
                   1827
Seq. No.
                   1114_7.R1040
Contig ID
                   LIB3107-003-Q1-K1-D6
5'-most EST
                   BLASTX
Method
                   q4160280
NCBI GI
BLAST score
                   251
                   2.0e-21
E value
Match length
                   88
% identity
                   (AJ006224) purple acid phosphatase [Ipomoea batatas]
NCBI Description
                   1828
Seq. No.
                   1116 1.R1040
Contig ID
                   pxt700941482.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3522938
BLAST score
                   887
                   2.0e-95
E value
                   330
Match length
                   57
% identity
                   (AC004411) unknown protein [Arabidopsis thaliana]
NCBI Description
                   1829
Seq. No.
                   1130 1.R1040
Contig ID
                   ASG32442DA-02-Q1-E1-H3
5'-most EST
                   BLASTX
Method
                   g1346882
NCBI GI
                   868
BLAST score
```

2.0e-93

E value



```
Match length 192
% identity 87
NCBI Description PHYTOENE SYNTHASE PRECURSOR (MEL5) >gi_2129505_pir__S56668
geranylgeranyl-diphosphate geranylgeranyltransferase (EC 2.5.1.32) precursor - muskmelon >gi_870985_emb_CAA85775_
(Z37543) phytoene synthase [Cucumis melo]

Seq. No. 1830
Contig ID 1130_2.R1040
5'-most EST fC-gmle700682749f1
Method BLASTN
NCBI GI 9433993
```

Method BLASTN
NCBI GI 9433993
BLAST score 51
E value 9.0e-20
Match length 95
% identity 88

NCBI Description C.annuum psyl mRNA for phytoene synthase

1831 Seq. No. 1133 1.R1040 Contig ID ncj700987183.h1 5'-most EST BLASTX Method q3913008 NCBI GI 1634 BLAST score 0.0e + 00E value 358 Match length % identity

NCBI Description FRUCTOSE-BISPHOSPHATE ALDOLASE, CYTOPLASMIC ISOZYME

>qi 3021338 emb CAA06308_ (AJ005041) cytosolic

fructose-1,6-bisphosphate aldolase [Cicer arietinum]

1832 Seq. No. 1133 2.R1040 Contig ID g4300477 5'-most EST BLASTN Method g3021337 NCBI GI 120 BLAST score 1.0e-60 E value 342 Match length 85 % identity

NCBI Description Cicer arietinum mRNA for cytosolic fructose-1,6-bisphosphate aldolase

 Seq. No.
 1833

 Contig ID
 1133_3.R1040

 5'-most EST
 jex700908902.h1

 Method
 BLASTN

 NCBI GI
 g3021337

NCBI GI g302133 BLAST score 119 E value 4.0e-60 Match length 267 % identity 87

NCBI Description Cicer arietinum mRNA for cytosolic fructose-1,6-bisphosphate aldolase

Seq. No. 1834 Contig ID 1133_6.R1040



5'-most EST LIB3167-050-P1-K1-A7

Seq. No.

1835

Contig ID

1135 1.R1040

5'-most EST

LIB3170-087-Q1-K1-A3

Method NCBI GI BLASTX q3036810

BLAST score E value

381

Match length

2.0e-36 133

% identity

61

NCBI Description

(AL022373) putative Myc-type transcription factor

[Arabidopsis thaliana]

Seq. No.

1836

Contig ID

1142 1.R1040

5'-most EST Method

LIB3109-047-Q1-K1-E2

NCBI GI BLAST score BLASTX q1351838

E value

2125 0.0e + 00

Match length

432

% identity

94

NCBI Description ACETYL-COENZYME A CARBOXYLASE CARBOXYL TRANSFERASE SUBUNIT

BETA >gi 984309 (U26948) beta-carboxyltransferase subunit

[Glycine max]

Seq. No.

1837

Contig ID

1142 2.R1040

5'-most EST

LIB3092-060-Q1-K1-G6

Method NCBI GI BLASTN g984307

BLAST score E value

1556 0.0e+00

Match length

1664 98

% identity NCBI Description

Glycine max ribosomal protein S16 (rps16) gene, partial cds, beta-carboxyltransferase (accD), photosystem I component (psaI), ORF 202 protein (ORF 203), ORF 151 protein (ORF 151), ORF 103 protein (ORF 103), ORF 229

precurso

Seq. No.

1838

Contig ID

1142 3.R1040

5'-most EST

jC-gmle01810034g10a2

Method NCBI GI BLASTN g984307

BLAST score

324

E value Match length 0.0e+00383

% identity

NCBI Description

Glycine max ribosomal protein S16 (rps16) gene, partial cds, beta-carboxyltransferase (accD), photosystem ${\tt I}$

component (psaI), ORF 202 protein (ORF 203), ORF 151 protein (ORF 151), ORF 103 protein (ORF 103), ORF 229

precurso



Seq. No. 1839

Contig ID 1142_7.R1040

5'-most EST LIB3093-024-Q1-K1-C1

Method BLASTN
NCBI GI g984307
BLAST score 230
E value 1.0e-126
Match length 348
% identity 93

NCBI Description Glycine max ribosomal protein S16 (rps16) gene, partial

cds, beta-carboxyltransferase (accD), photosystem I component (psaI), ORF 202 protein (ORF 203), ORF 151 protein (ORF 151), ORF 103 protein (ORF 103), ORF 229

precurso

Seq. No. 1840

Contig ID 1144 1.R1040

5'-most EST LIB3139-067-P1-N1-E5

Method BLASTX
NCBI GI g3402716
BLAST score 299
E value 1.0e-26
Match length 233
% identity 34

NCBI Description (AC004261) unknown protein [Arabidopsis thaliana]

Seq. No. 1841

Contig ID 1149 1.R1040

5'-most EST LIB3\overline{1}39-008-P1-N1-C9

Method BLASTN
NCBI GI g19500
BLAST score 164
E value 1.0e-86
Match length 433
% identity 88

NCBI Description L.polyphyllus mRNA for pPLZ12 protein

Seq. No. 1842

Contig ID 1149_2.R1040

5'-most EST LIB3074-031-Q1-K1-B9

Seq. No. 1843

Contig ID 1155_1.R1040 5'-most EST gsv701046559.h1

Method BLASTN
NCBI GI g396818
BLAST score 222
E value 1.0e-121
Match length 734
% identity 83

NCBI Description P.sativum pspor1 mRNA encoding porin

Seq. No. 1844

Contig ID 1157_1.R1040

5'-most EST uC-gmrominsoy256g12b1

Method BLASTN NCBI GI g4115332



BLAST score 444
E value 0.0e+00
Match length 879
% identity 21

NCBI Description Pisum sativum (Alaska) ubiquitin (PUB1) gene, complete cds

Seq. No. 1845

Contig ID 1157 2.R1040

5'-most EST LIB3093-033-Q1-K1-F8

Method BLASTX
NCBI GI g170352
BLAST score 1835
E value 0.0e+00
Match length 368
% identity 35

NCBI Description (M74101) hexameric polyubiquitin [Nicotiana sylvestris] >gi 870792 (L05361) polyubiquitin [Arabidopsis thaliana]

>gi_4115333 (L81139) ubiquitin [Pisum sativum] >gi_4115335

(L8 $\overline{1}$ 140) ubiquitin [Pisum sativum]

Seq. No. 1846

Contig ID 1157 3.R1040

5'-most EST LIB3050-011-Q1-E1-D11

Method BLASTN
NCBI GI g4115334
BLAST score 340
E value 0.0e+00
Match length 620
% identity 20

NCBI Description Pisum sativum (Alaska) ubiquitin (PUB2) gene, complete cds

Seq. No. 1847

Contig ID 1157_4.R1040

5'-most EST LIB3138-025-Q1-N1-A2

Method BLASTN
NCBI GI g4115338
BLAST score 206
E value 1.0e-112
Match length 337
% identity 22

NCBI Description Pisum sativum (Alaska) ubuquitin (PUB4) gene, complete cds

Seq. No. 1848

Contig ID 1158 1.R1040

5'-most EST LIB3107-062-Q1-K1-A5

Method BLASTN
NCBI GI g12178
BLAST score 511
E value 0.0e+00
Match length 745
% identity 92

NCBI Description Pea chloroplast rpl36 (partial), rps11, rpoA, and petD

(partial) genes for ribosomal proteins L36 and S11, RNA polymerase alpha subunit, and cytochrome bf complex subunit

ĪV

Seq. No. 1849



1158 3.R1040 Contig ID wvk700686173.h1 5'-most EST BLASTN Method g2924257 NCBI GI BLAST score 151 E value 1.0e-79 203 Match length 94 % identity NCBI Description Tobacco chloroplast genome DNA Seq. No. 1850 1162 1.R1040 Contig ID LIB3040-050-Q1-E1-C5 5'-most EST BLASTN Method g2264309 NCBI GI BLAST score 53 E value 2.0e-20 Match length 472 82 % identity Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description MJJ3, complete sequence [Arabidopsis thaliana] 1851 Seq. No. 1164 1.R1040 Contig ID jC-gmf102220096g11a1 5'-most EST BLASTX Method g461812 NCBI GI BLAST score 342 E value 5.0e-32 162 Match length % identity 40 CYTOCHROME P450 72 (CYPLXXII) (PROBABLE NCBI Description GERANIOL-10-HYDROXYLASE) (GE10H) >gi_167484 (L10081) Cytochrome P-450 protein [Catharanthus roseus] >gi 445604_prf__1909351A cytochrome P450 [Catharanthus roseus] 1852 Seq. No. 1167 1.R1040 Contig ID uC-gmropic032h12b1 5'-most EST BLASTN Method q470126 NCBI GI 220 BLAST score E value 1.0e-120 536 Match length % identity 85 NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19 1853 Seq. No. 1167 2.R1040 Contig ID 5'-most EST LIB3106-051-Q1-K1-H5

Method BLASTN
NCBI GI g470126
BLAST score 170
E value 3.0e-90
Match length 566
% identity 83



NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19

Seq. No. 1854

Contig ID 1167_3.R1040

5'-most EST uC-gmropic010a09b1

Method BLASTN
NCBI GI 9470126
BLAST score 105
E value 1.0e-51
Match length 333
% identity 83

NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19

Seq. No. 1855

Contig ID 1167_4.R1040 5'-most EST jsh701064147.h1

Method BLASTX
NCBI GI g3377797
BLAST score 417
E value 4.0e-79
Match length 174
% identity 89

NCBI Description (AF075597) Similar to 60S ribosome protein L19; coded for

by A. thaliana cDNA T04719; coded for by A. thaliana cDNA H36046; coded for by A. thaliana cDNA T44067; coded for by A. thaliana cDNA T14056; coded for by A. thaliana cDNA

R90691 [Ara

Seq. No. 1856

1167 5.R1040 Contig ID g5688162 5'-most EST BLASTN Method g169928 NCBI GI BLAST score 81 2.0e-37 E value 89 Match length 98 % identity

NCBI Description Glycine max alpha'-type beta conglycinin storage protein

gene, complete cds, clone ch4A

Seq. No. 1857

Contig ID 1167_6.R1040

5'-most EST LIB3170-035-Q1-K1-G2

Method BLASTN
NCBI GI 9470126
BLAST score 94
E value 2.0e-45
Match length 242
% identity 85

NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19

Seq. No. 1858

Contig ID 1167 7.R1040

5'-most EST LIB3050-011-Q1-E1-F12

Method BLASTN NCBI GI g470126 BLAST score 90

NCBI GI BLAST score

E value

Match length

378

2.0e-45 123



```
5.0e-43
E value
                  238
Match length
% identity
NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19
                  1859
Seq. No.
                  1175 1.R1040
Contig ID
                  LIB3107-038-Q1-K1-A3
5'-most EST
                  BLASTX
Method
                  g3378650
NCBI GI
BLAST score
                  1274
                  1.0e-141
E value
                   313
Match length
                   76
% identity
                  (X97606) abscisic acid activated [Medicago sativa]
NCBI Description
Seq. No.
                   1860
                   1175 2.R1040
Contig ID
                   rca701000823.h1
5'-most EST
Method
                   BLASTX
                   g3378650
NCBI GI
BLAST score
                   490
                   3.0e-49
E value
                   113
Match length
                   80
% identity
                  (X97606) abscisic acid activated [Medicago sativa]
NCBI Description
                   1861
Seq. No.
                   1175 3.R1040
Contig ID
                   LIB3093-009-Q1-K1-G11
5'-most EST
                   BLASTN
Method
                   g3378649
NCBI GI
                   150
BLAST score
                   1.0e-78
E value
                   314
Match length
                   87
% identity
NCBI Description M.sativa mRNA translated from abscisic activated gene
                   1862
Seq. No.
                   1175 4.R1040
Contig ID
                   uC-gmronoir023c03b1
5'-most EST
                   BLASTN
Method
                   g3378649
NCBI GI
                   155
BLAST score
                   1.0e-81
E value
Match length
                   311
                   87
 % identity
NCBI Description M.sativa mRNA translated from abscisic activated gene
                   1863
 Seq. No.
                   1175 5.R1040
 Contig ID
                   zsg701120171.hl
 5'-most EST
                   BLASTX
 Method
                   g3378650
```



% identity NCBI Description (X97606) abscisic acid activated [Medicago sativa]

1864 Seq. No.

1175 6.R1040 Contig ID

 $LIB3\overline{0}92-022-Q1-K1-G11$ 5'-most EST

BLASTN Method g3378649 NCBI GI 103 BLAST score 1.0e-50 E value 278 Match length 84 % identity

NCBI Description M.sativa mRNA translated from abscisic activated gene

1865 Seq. No.

Contig ID 1179 1.R1040

LIB3072-053-Q1-E1-F3 5'-most EST

Method BLASTX g2119749 NCBI GI 249 BLAST score 7.0e-21 E value 78 Match length 60 % identity

oleosin - almond NCBI Description

Seq. No. 1866

1181 1.R1040 Contig ID

uC-gmrominsoy313h12b1 5'-most EST

1867 Seq. No.

1182 1.R1040 Contig ID

LIB3027-010-Q1-B1-F2 5'-most EST

1868 Seq. No.

1184 1.R1040 Contig ID

LIB3027-010-Q1-B1-F4 5'-most EST

1869 Seq. No.

1187 1.R1040 Contig ID

LIB3139-081-P1-N1-C1 5'-most EST

BLASTX Method q2245004 NCBI GI 345 BLAST score E value 2.0e-32 Match length 106 % identity 59

(Z97341) similarity to membrane transport protein NCBI Description

[Arabidopsis thaliana]

1870 Seq. No.

1188 1.R1040 Contig ID 5'-most EST seb700653874.hl

1871 Seq. No.

1188 2.R1040 Contig ID

jC-gmst02400072b08a1 5'-most EST



```
Seq. No.
                  1872
Contig ID
                  1189 1.R1040
5'-most EST
                  jC-gmle01810012e01a1
Method
                  BLASTX
NCBI GI
                  q3928760
BLAST score
                  966
                  1.0e-104
E value
                  265
Match length
% identity
                  75
NCBI Description (AB011797) homolog to plastid-lipid-associated protein
                  [Citrus unshiu]
Seq. No.
                  1873
Contig ID
                  1189 2.R1040
5'-most EST
                  LIB3170-056-Q1-K1-A4
Method
                  BLASTX
                  g3928760
NCBI GI
BLAST score
                  356
                  1.0e-33
E value
Match length
                  113
% identity
                  66
                 (AB011797) homolog to plastid-lipid-associated protein
NCBI Description
                  [Citrus unshiu]
Seq. No.
                  1874
Contig ID
                  1189 3.R1040
5'-most EST
                  LIB3106-085-Q1-K1-D3
Seq. No.
                  1875
Contig ID
                  1189 4.R1040
5'-most EST
                  V4R-02-Q1-B1-H1
Seq. No.
                  1876
                  1195 1.R1040
Contig ID
5'-most EST
                  fua701041821.h1
                  BLASTX
Method
NCBI GI
                  q3355658
BLAST score
                  192
                  1.0e-14
E value
Match length
                  124
                  38
% identity
NCBI Description (AJ009691) SMC2orf [Podocoryne carnea]
Seq. No.
                  1877
Contig ID
                  1199 1.R1040
5'-most EST
                  jC-qmst02400023e02a1
Method
                  BLASTX
```

NCBI GI g1705574 BLAST score 176 3.0e-12 E value Match length 172 % identity 33

NCBI Description CALPAIN (CALCIUM-ACTIVATED NEUTRAL PROTEINASE) (CANP)

>gi_562289_emb_CAA55298_ (X78555) calpain [Drosophila

melanogaster]

1878

Seq. No.



Contig ID 1202_1.R1040 5'-most EST zpv700758807.h1

Method BLASTX
NCBI GI g633890
BLAST score 455
E value 2.0e-45
Match length 143
% identity 61

NCBI Description (S72926) glucose and ribitol dehydrogenase homolog [Hordeum

vulgare]

Seq. No. 1879

Contig ID 1203 1.R1040

5'-most EST jC-gmf102220061g03a1

Method BLASTX
NCBI GI g4139266
BLAST score 937
E value 1.0e-101
Match length 325
% identity 60

NCBI Description (AF112444) L-asparaginase [Lupinus luteus]

Seq. No. 1880

Contig ID 1203 2.R1040

5'-most EST uC-gmrominsoy129c01b1

Method BLASTX
NCBI GI g2462826
BLAST score 219
E value 2.0e-17
Match length 137
% identity 40

NCBI Description (AF000657) unknown protein [Arabidopsis thaliana]

Seq. No. 1881

Contig ID 1203_3.R1040

5'-most EST LIB3049-001-Q1-E1-C8

Seq. No. 1882

Contig ID 1203_4.R1040

5'-most EST jC-gmro02910005f01a1

Method BLASTX
NCBI GI g2462826
BLAST score 173
E value 3.0e-12
Match length 81
% identity 44

NCBI Description (AF000657) unknown protein [Arabidopsis thaliana]

Seq. No. 1883

Contig ID 1203 6.R1040

5'-most EST jC-gmf102220061g03d1

Seq. No. 1884

Contig ID 1203 7.R1040

5'-most EST uC-gmflminsoy044h10b1

Method BLASTX NCBI GI g1703446



BLAST score 235 E value 1.0e-19 Match length 83 % identity 61

NCBI Description L-ASPARAGINASE (L-ASPARAGINE AMIDOHYDROLASE)

>gi_1076292_pir__S53127 asparaginase - Arabidopsis thaliana >gi_735918_emb_CAA84367_ (Z34884) asparaginase [Arabidopsis

thaliana]

Seq. No. 1885

Contig ID 1203_11.R1040 5'-most EST zhf700961256.h1

Method BLASTX
NCBI GI g1703446
BLAST score 186
E value 3.0e-14
Match length 65
% identity 63

NCBI Description L-ASPARAGINASE (L-ASPARAGINE AMIDOHYDROLASE)

>gi_1076292_pir__S53127 asparaginase - Arabidopsis thaliana >gi_735918_emb_CAA84367_ (Z34884) asparaginase [Arabidopsis

thaliana]

Seq. No. 1886

Contig ID 1204_1.R1040 5'-most EST awf700839244.h1

Method BLASTX
NCBI GI g4204308
BLAST score 237
E value 7.0e-20
Match length 43
% identity 91

NCBI Description (AC003027) lcl_prt_seq No definition line found

[Arabidopsis thaliana]

Seq. No. 1887

Contig ID 1206_1.R1040

5'-most EST LIB3027-010-Q1-B1-C9

Seq. No. 1888

Contig ID 1210 1.R1040

5'-most EST uC-gmflminsoy005a10b1

Method BLASTX
NCBI GI g2623309
BLAST score 702
E value 2.0e-73
Match length 435
% identity 42

NCBI Description (AC002409) similar to tgacg-specific DNA-binding protein

[Arabidopsis thaliana]

Seq. No. 1889

Contig ID 1212 1.R1040

5'-most EST LIB3029-006-Q1-B1-H7

Method BLASTX NCBI GI g1710077 BLAST score 558



```
2.0e-74
E value
                      190
Match length
                      71
% identity
                      PEROXIREDOXIN (REHYDRIN HOMOLOG) (B15C)
NCBI Description
                      >gi_2130029_pir__S60285 B15C protein - barley
                      >gi_471321 emb_CAA54066 (X76605) HvB15C [Hordeum vulgare]
>gi_1694833_emb_CAA65387 (X96551) peroxiredoxin [Hordeum
                      vulgare]
                      1890
Seq. No.
                      1215 1.R1040
Contig ID
                      LIB3040-022-Q1-E1-D9
5'-most EST
```

BLASTX Method g3759184 NCBI GI BLAST score 805 E value 4.0e-86 Match length 225 % identity 68

(AB018441) phi-1 [Nicotiana tabacum] NCBI Description

Seq. No. 1219 1.R1040 Contig ID LIB3051-097-Q1-K1-G7 5'-most EST BLASTN Method g22075 NCBI GI BLAST score 167 9.0e-89

E value Match length 247 % identity 92

NCBI Description Vigna unguiculata cDNA for stored cotyledon mRNA

1892 Seq. No.

1223 1.R1040 Contig ID

jC-gmle01810030e08a2 5'-most EST

1891

BLASTX Method g1173555 NCBI GI 1661 BLAST score 0.0e + 00E value 350 Match length % identity

(U31544) UDP-galactose-4-epimerase [Pisum sativum] NCBI Description

1893 Seq. No.

1223 2.R1040 Contig ID jex700910050.h1 5'-most EST

BLASTX Method q1173555 NCBI GI 180 BLAST score E value 1.0e-13 43 Match length % identity

(U31544) UDP-galactose-4-epimerase [Pisum sativum] NCBI Description

1894 Seq. No.

1223 3.R1040 Contig ID wrq700787631.h2 5'-most EST

BLASTX Method



```
g1173555
NCBI GI
                  232
BLAST score
                  1.0e-19
E value
Match length
                  71
% identity
                   69
                   (U31544) UDP-galactose-4-epimerase [Pisum sativum]
NCBI Description
                  1895
Seq. No.
                   1225 1.R1040
Contig ID
                   uC-gmrominsoy101e03b1
5'-most EST
                   BLASTX
Method
                   g2244814
NCBI GI
                   722
BLAST score
                   2.0e-76
E value
Match length
                   189
                   75
% identity
                   (Z97336) protein kinase [Arabidopsis thaliana]
NCBI Description
                   1896
Seq. No.
                   1228 1.R1040
Contig ID
                   LIB3109-054-Q1-K1-H11
5'-most EST
                   BLASTX
Method
                   g4510348
NCBI GI
                   774
BLAST score
                   2.0e-82
E value
                   213
Match length
                   64
% identity
                   (AC006921) unknown protein [Arabidopsis thaliana]
NCBI Description
                   1897
Seq. No.
                   1230 1.R1040
Contig ID
                   LIB3027-009-Q1-B1-E6
5'-most EST
                   BLASTX
Method
                   g3892055
NCBI GI
BLAST score
                   325
                   3.0e-30
E value
                   121
Match length
% identity
                   55
                   (AC002330) putative transport protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   1898
                   1243 1.R1040
Contig ID
                   LIB3049-039-Q1-E1-A2
5'-most EST
Method
                   BLASTX
                   q3650033
NCBI GI
BLAST score
                   959
                   1.0e-104
E value
Match length
                   302
                   66
% identity
                   (AC005396) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1899
                   1243_2.R1040
Contig ID
                   uC-gmropic036g10b1
 5'-most EST
```

426

BLASTX

g3650033

Method

NCBI GI



BLAST score 328 E value 2.0e-30 Match length 85 % identity 78

NCBI Description (AC005396) unknown protein [Arabidopsis thaliana]

Seq. No. 1900

Contig ID 1244_1.R1040

5'-most EST LIB3027-009-Q1-B1-B12

Seq. No. 1901

Contig ID 1246_1.R1040

5'-most EST LIB3029-007-Q1-B1-A2

Method BLASTN
NCBI GI g256634
BLAST score 810
E value 0.0e+00
Match length 872
% identity 45

NCBI Description KTi1=Kunitz trypsin inhibitor KTi1, KTi2=Kunitz trypsin

inhibitor KTi2 [soybeans, Genomic, 3269 nt]

Seq. No. 1902

Contig ID 1246_3.R1040 5'-most EST vwf700673594.h1

Method BLASTN
NCBI GI g256634
BLAST score 185
E value 1.0e-100
Match length 250
% identity 50

NCBI Description KTi1=Kunitz trypsin inhibitor KTi1, KTi2=Kunitz trypsin

inhibitor KTi2 [soybeans, Genomic, 3269 nt]

Seq. No. 1903

Contig ID 1247_1.R1040

5'-most EST uC-gmflminsoy042f10b1

Method BLASTN
NCBI GI g2970653
BLAST score 458
E value 0.0e+00
Match length 738
% identity 92

NCBI Description Vigna unguiculata ferritin subunit cowpea2 precursor, mRNA,

nuclear gene encoding chloroplast protein, complete cds

Seq. No. 1904

Contig ID 1247_2.R1040

5'-most EST LIB3049-045-Q1-E1-C6

Method BLASTN
NCBI GI g2970653
BLAST score 198
E value 1.0e-107
Match length 346
% identity 92

NCBI Description Vigna unguiculata ferritin subunit cowpea2 precursor, mRNA,

nuclear gene encoding chloroplast protein, complete cds



Seq. No. 1905

Contig ID 1248 1.R1040

5'-most EST LIB3139-026-P1-N1-C10

Method BLASTX
NCBI GI g3023832
BLAST score 1513
E value 1.0e-169
Match length 320
% identity 52

NCBI Description GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT 1

>gi 1835161 emb_CAB06618_ (Z84820) G protein beta subunit

[Nicotiana tabacum]

Seq. No. 1906

Contig ID 1248_2.R1040

5'-most EST uC-gmrominsoy167g11b1

Method BLASTX
NCBI GI g3023832
BLAST score 663
E value 1.0e-69
Match length 142
% identity 88

NCBI Description GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT 1

>qi 1835161 emb CAB06618 (Z84820) G protein beta subunit

[Nicotiana tabacum]

Seq. No. 1907

Contig ID 1248_3.R1040

5'-most EST LIB3109-021-Q1-K2-H8

Method BLASTX
NCBI GI g3023841
BLAST score 304
E value 1.0e-27
Match length 68
% identity 87

NCBI Description GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT

>gi 1695179_emb_CAA70704_ (Y09513) G protein beta subunit

[Nicotiana plumbaginifolia]

Seq. No. 1908

Contig ID 1248_4.R1040 5'-most EST zpv700762067.h1

Method BLASTX
NCBI GI g3023841
BLAST score 254
E value 4.0e-22
Match length 70
% identity 69

NCBI Description GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT

>gi 1695179_emb_CAA70704_ (Y09513) G protein beta subunit

[Nicotiana plumbaginifolia]

Seq. No. 1909

Contig ID 1250_1.R1040 5'-most EST g4305784



```
1910
Seq. No.
                  1253 1.R1040
Contig ID
                  gsv701046538.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2506470
                  2270
BLAST score
                  0.0e + 00
E value
                  541
Match length
% identity
                  ALPHA-1,4 GLUCAN PHOSPHORYLASE, L ISOFORM PRECURSOR (STARCH
NCBI Description
                  PHOSPHORYLASE L) >gi_1616637_emb_CAA85354_ (Z36880)
                   alpha-1,4 Glucan Phosphorylase, L isoform precursor [Vicia
                   1911
Seq. No.
Contig ID
                   1253 2.R1040
                   jC-gmst02400070b12a1
5'-most EST
Method
                   BLASTN
                   q534971
NCBI GI
BLAST score
                   122
                   6.0e-62
E value
Match length
                   178
                   92
% identity
                  V.faba (var. minor) mRNA for alpha 1,4-glucan phosphorylase
NCBI Description
                   L isoform
                   1912
Seq. No.
Contig ID
                   1253 3.R1040
                   LIB3074-036-Q1-K1-C4
5'-most EST
                   BLASTN
Method
                   q11576
NCBI GI
BLAST score
                   583
                   0.0e + 00
E value
                   1120
Match length
                   99
% identity
                   Soybean plastid DNA for rps12, rps7, 16S rRNA, tRNA-Val,
NCBI Description
                   NADH dehydrogenase and ORF
                   1913
Seq. No.
                   1253 8.R1040
Contig ID
                   LIB3093-053-Q1-K1-E4
5'-most EST
                   BLASTN
Method
                   g11576
NCBI GI
                   116
BLAST score
                   2.0e-58
E value
                   413
Match length
                   93
 % identity
                   Soybean plastid DNA for rps12, rps7, 16S rRNA, tRNA-Val,
NCBI Description
                   NADH dehydrogenase and ORF
                   1914
 Seq. No.
```

Contig ID 1253_9.R1040 5'-most EST leu701144377.h1

Method BLASTN
NCBI GI g11576
BLAST score 111
E value 1.0e-55



```
Match length 170
% identity 93
```

NCBI Description Soybean plastid DNA for rps12, rps7, 16S rRNA, tRNA-Val,

NADH dehydrogenase and ORF

Seq. No. 1915

Contig ID 1267_1.R1040 5'-most EST awf700842486.h1

Method BLASTN
NCBI GI g3046856
BLAST score 35
E value 1.0e-09
Match length 301
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MXI22, complete sequence [Arabidopsis thaliana]

Seq. No. 1916

Contig ID 1272 1.R1040 5'-most EST g5688109 Method BLASTN g170005 NCBI GI BLAST score 1008 0.0e+00E value 1022 Match length % identity 100

NCBI Description Soybean lectin (Le1) gene, complete cds

Seq. No. 1917

Contig ID 1272 2.R1040

5'-most EST LIB3092-053-Q1-K1-D3

Seq. No. 1918

Contig ID 1273_1.R1040

5'-most EST LIB3107-061-Q1-K1-E4

Method BLASTX
NCBI GI g4115934
BLAST score 593
E value 3.0e-61
Match length 128
% identity 79

NCBI Description (AF118223) contains similarity to Methanobacterium

thermoautotrophicum transcriptional regulator (GB:AE000850)

[Arabidopsis thaliana]

Seq. No. 1919

Contig ID 1273_2.R1040

5'-most EST LIB3107-016-Q1-K1-D1

Method BLASTX
NCBI GI g4115939
BLAST score 237
E value 7.0e-20
Match length 57
% identity 40

NCBI Description (AF118223) contains similarity to Methanobacterium

thermoautotrophicum transcriptional regulator (GB:AE000850)

[Arabidopsis thaliana]



```
1920
Seq. No.
                  1274_1.R1040
Contig ID
                  LIB3109-001-Q1-K2-B7
5'-most EST
                  BLASTX
Method
                  q401188
NCBI GI
                   383
BLAST score
                   4.0e-48
E value
                   305
Match length
% identity
                   22
                  NUCLEOLYSIN TIA-1 >gi_107354_pir__A39293 cytotoxic
NCBI Description
                   granule-associated RNA-binding protein TIA1 precursor,
                   leukocyte - human
Seq. No.
                   1921
                   1274 2.R1040
Contig ID
                   LIB3051-001-Q1-E1-D4
5'-most EST
                   1922
Seq. No.
                   1274 3.R1040
Contig ID
                   LIB3170-042-Q1-K1-C2
5'-most EST
                   1923
Seq. No.
                   1276 1.R1040
Contig ID
                   LIB3107-075-Q1-K1-E5
5'-most EST
                   BLASTX
Method
                   g3193298
NCBI GI
                   190
BLAST score
                   6.0e-14
E value
Match length
                   108
                   35
% identity
                   (AF069298) T14P8.17 gene product [Arabidopsis thaliana]
NCBI Description
                   1924
Seq. No.
                   1279 1.R1040
Contig ID
5'-most EST
                   LIB3106-070-P1-K1-F1
Method
                   BLASTN
                   q3603398
NCBI GI
                   336
BLAST score
E value
                   0.0e + 00
                   1019
Match length
                   87
% identity
                   Medicago sativa cinnamyl-alcohol dehydrogenase (MsaCad2)
NCBI Description
                   mRNA, complete cds
                   1925
Seq. No.
                   1279 2.R1040
Contig ID
                   pmv700890767.h1
 5'-most EST
Method
                   BLASTN
                   g19594
NCBI GI
```

Method BLASTN
NCBI GI g19594
BLAST score 143
E value 2.0e-74
Match length 338
% identity 88

NCBI Description M.sativa encoding cinnamyl alcohol dehydrogenase

Seq. No. 1926



```
1279 3.R1040
Contig ID
                  LIB3138-062-Q1-N1-E10
5'-most EST
Method
                  BLASTN
NCBI GI
                  g3603398
BLAST score
                  140
                  2.0e-72
E value
                  387
Match length
                  89
% identity
                  Medicago sativa cinnamyl-alcohol dehydrogenase (MsaCad2)
NCBI Description
                  mRNA, complete cds
                  1927
Seq. No.
                  1279 5.R1040
Contig ID
                  sat701009719.h1
5'-most EST
                  BLASTN
Method
                  g3603398
NCBI GI
BLAST score
                  129
                   3.0e-66
E value
Match length
                   284
                   86
% identity
                  Medicago sativa cinnamyl-alcohol dehydrogenase (MsaCad2)
NCBI Description
                   mRNA, complete cds
                   1928
Seq. No.
                   1281 1.R1040
Contig ID
                   jC-gmro02800043h04a1
5'-most EST
                   BLASTX
Method
                   g2865175
NCBI GI
BLAST score
                   412
                   3.0e-40
E value
                   132
Match length
                   61
% identity
                   (AB010945) AtRer1A [Arabidopsis thaliana]
NCBI Description
                   1929
Seq. No.
                   1281_2.R1040
Contig ID
                   pxt700941249.h1
5'-most EST
                   BLASTX
Method
                   q2642434
NCBI GI
                   229
BLAST score
                   5.0e-19
E value
                   102
Match length
                   46
% identity
                   (AC002391) putative Rerl protein [Arabidopsis thaliana]
NCBI Description
                   1930
Seq. No.
                   1282 1.R1040
Contig ID
                   leu701152518.hl
 5'-most EST
Method
                   BLASTN
                   q343480
NCBI GI
BLAST score
                   877
                   0.0e+00
E value
                   1323
Match length
```

tobacco chloroplast atpase gene (b and e subunits) and

92

flanks

% identity

NCBI Description



```
1931
Seq. No.
Contig ID
5'-most EST
```

1282 6.R1040 fde700872028.hl

Method BLASTN g3929530 NCBI GI 168 BLAST score 1.0e-89 E value 244 Match length 96

% identity Anisoptera marginata ATP synthase beta subunit (atpB) gene, NCBI Description chloroplast gene encoding chloroplast protein, partial cds

1932 Seq. No.

1283 1.R1040 Contig ID 5'-most EST sat701003615.h1

Method BLASTN g3043427 NCBI GI 278 BLAST score 1.0e-155 E value 594 Match length % identity 87

NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S5

1933 Seq. No.

1283 2.R1040 Contig ID rca700999874.hl 5'-most EST

1934 Seq. No.

1283 3.R1040 Contig ID

uC-gmrominsoy068d03b1 5'-most EST

BLASTN Method g3043427 NCBI GI 248 BLAST score 1.0e-137 E value Match length 592

85 % identity

NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S5

1935 Seq. No.

1283 4.R1040 Contig ID

LIB3170-014-Q1-K1-E1 5'-most EST

BLASTN Method q3043427 NCBI GI 183 BLAST score E value 4.0e-98 567 Match length % identity

NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S5

1936 Seq. No.

1283 6.R1040 Contig ID

LIB3051-015-Q1-E1-C1 5'-most EST

BLASTN Method g3043427 NCBI GI BLAST score 117 4.0e-59 E value 279 Match length



% identity 87 NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S5

Seq. No. 1937

Contig ID 1285_1.R1040

5'-most EST LIB3049-048-Q1-E1-A5

Method BLASTN
NCBI GI g20657
BLAST score 362
E value 0.0e+00
Match length 753
% identity 89

NCBI Description P.sativum Cab II gene for chlorophyll a/b-binding protein

Seq. No. 1938

Contig ID 1285 2.R1040

5'-most EST LIB3170-023-Q1-J1-B4

Method BLASTX
NCBI GI g3560529
BLAST score 194
E value 7.0e-15
Match length 49
% identity 80

NCBI Description (AF039598) light harvesting chlorophyll A/B binding protein

[Prunus persica]

Seq. No. 1939

Contig ID 1285_3.R1040 5'-most EST hyd700727351.h1

Method BLASTN
NCBI GI g20657
BLAST score 229
E value 1.0e-126
Match length 450
% identity 91

NCBI Description P.sativum Cab II gene for chlorophyll a/b-binding protein

Seq. No. 1940

Contig ID 1287_1.R1040

5'-most EST LIB3106-019-Q1-K1-H2

Method BLASTX
NCBI GI g3183247
BLAST score 1178
E value 1.0e-129
Match length 391
% identity 58

NCBI Description PUTATIVE GTP-BINDING PROTEIN W08E3.3

>gi_3880615_emb_CAB07131_ (Z92773) predicted using Genefinder; Similarity to Yeast hypothetical 44.2 KD

protein, putative GTP-binding protein (SW:P38219); cDNA EST

EMBL:D64516 comes from this gene; cDNA EST EMBL:D65777

comes from this gene; cDNA EST EMB

Seq. No. 1941

Contig ID 1290_1.R1040 5'-most EST kmv700739112.h1

Method BLASTX

NCBI Description



```
q4056479
NCBI GI
BLAST score
                  291
                  4.0e-26
E value
                  82
Match length
% identity
                  (AC005896) unknown protein [Arabidopsis thaliana]
NCBI Description
                  1942
Seq. No.
                  1291 1.R1040
Contig ID
5'-most EST
                  LIB3027-008-Q1-B1-A3
Seq. No.
                   1292 1.R1040
Contig ID
                   LIB3029-012-Q1-B1-B1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1839188
BLAST score
                   539
E value
                   7.0e-55
                   157
Match length
% identity
                  (U86081) root hair defective 3 [Arabidopsis thaliana]
NCBI Description
                   1944
Seq. No.
                   1293 1.R1040
Contig ID
                   leu701153506.hl
5'-most EST
Seq. No.
                   1945
                   1294 1.R1040
Contig ID
                   q5687992
5'-most EST
Method
                   BLASTX
                   g729470
NCBI GI
BLAST score
                   1608
                   1.0e-180
E value
                   356
Match length
                   83
% identity
                   MITOCHONDRIAL FORMATE DEHYDROGENASE PRECURSOR
NCBI Description
                   (NAD-DEPENDENT FORMATE DEHYDROGENASE) (FDH)
                   >gi 542089_pir__JQ2272 formate dehydrogenase (EC 1.2.1.2)
                   precursor, mitochondrial - potato >gi_297798_emb_CAA79702_
                   (Z21493) mitochondrial formate dehydrogenase precursor
                   [Solanum tuberosum]
                   1946
 Seq. No.
                   1294 2.R1040
 Contig ID
                   kl1701213327.h1
 5'-most EST
                   BLASTX
Method
                   g729470
 NCBI GI
                   518
 BLAST score
                   1.0e-52
 E value
                   115
 Match length
                   85
 % identity
```

435

[Solanum tuberosum]

MITOCHONDRIAL FORMATE DEHYDROGENASE PRECURSOR

>gi_542089_pir__JQ2272 formate dehydrogenase (EC 1.2.1.2)
precursor, mitochondrial - potato >gi_297798_emb_CAA79702_
(Z21493) mitochondrial formate dehydrogenase precursor

(NAD-DEPENDENT FORMATE DEHYDROGENASE) (FDH)



1947 Seq. No. Contig ID 1294 4.R1040 hrw701059575.hl 5'-most EST BLASTX

Method g729470 NCBI GI 536 BLAST score 7.0e-55 E value 139 Match length 77 % identity

MITOCHONDRIAL FORMATE DEHYDROGENASE PRECURSOR NCBI Description (NAD-DEPENDENT FORMATE DEHYDROGENASE) (FDH)

>gi_542089_pir__JQ2272 formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato >gi_297798_emb_CAA79702_ (Z21493) mitochondrial formate dehydrogenase precursor

[Solanum tuberosum]

1948 Seq. No.

1294 5.R1040 Contig ID

LIB3072-059-Q1-K1-B5 5'-most EST

BLASTX Method g729470 NCBI GI 293 BLAST score 2.0e-26 E value 71 Match length 75 % identity

MITOCHONDRIAL FORMATE DEHYDROGENASE PRECURSOR NCBI Description (NAD-DEPENDENT FORMATE DEHYDROGENASE) (FDH)

>gi_542089_pir__JQ2272 formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato >gi_297798_emb_CAA79702_

(Z21493) mitochondrial formate dehydrogenase precursor

[Solanum tuberosum]

1949 Seq. No.

1294 7.R1040 Contig ID

5'-most EST LIB3072-017-Q1-E1-H12

1950 Seq. No.

1294 8.R1040 Contig ID 5'-most EST rca701002059.h1

1951 Seq. No.

1294 9.R1040 Contig ID

jC-qmf102220063f11d1 5'-most EST

Method BLASTX q4062934 NCBI GI BLAST score 444 4.0e-44 E value Match length 85 91 % identity

(D88272) formate dehydrogenase [Hordeum vulgare] NCBI Description

1952 Seq. No.

1294 11.R1040 Contig ID kmv700742307.h1 5'-most EST

BLASTX Method g3821730 NCBI GI



```
BLAST score
                  142
                  3.0e-09
E value
                  55
Match length
% identity
NCBI Description (Z99991) formate dehydrogenase [Solanum tuberosum]
                  1953
Seq. No.
                  1295 1.R1040
Contig ID
                  LIB3109-029-Q1-K1-H1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2245086
                  540
BLAST score
                   4.0e-55
E value
                  158
Match length
% identity
NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   1295 2.R1040
Contig ID
                   jC-gmst02400016b09d1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2245086
                   329
BLAST score
                   2.0e-30
E value
                   89
Match length
                   75
% identity
NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]
                   1955
Seq. No.
                   1296 1.R1040
Contig ID
                   ssr700555601.hl
5'-most EST
Method
                   BLASTX
                   q231660
NCBI GI
                   1146
BLAST score
                   1.0e-125
E value
                   636
Match length
% identity
                   44
NCBI Description HYPOTHETICAL 226 KD PROTEIN (ORF 1901)
                   1956
Seq. No.
                   1296 2.R1040
Contig ID
                   LIB3138-059-Q1-N1-C5
 5'-most EST
                   BLASTX
Method
                   g82200
NCBI GI
                   734
BLAST score
                   2.0e-77
 E value
                   404
Match length
                   46
 % identity
NCBI Description hypothetical protein 1244 - common tobacco chloroplast
 Seq. No.
                   1957
 Contig ID
                   1296 9.R1040
                   uC-gmropic097d09b1
 5'-most EST
                   BLASTN
 Method
                   g2924257
 NCBI GI
 BLAST score
                   55
```

4.0e-22

E value



Match length 203 % identity 82

NCBI Description Tobacco chloroplast genome DNA

Seq. No. Contig ID 1958

Contig ID 5'-most EST 1296_10.R1040 leu701148723.h1

Seq. No.

1959

Contig ID 5'-most EST

1298_1.R1040 rlr700896228.h1

Method BLASTN
NCBI GI g19701
BLAST score 206
E value 1.0e-111
Match length 470
% identity 86

NCBI Description N.plumbaginifolia mRNA NeIF-5A2 for initiation factor 5A(2)

Seq. No.

1960

Contig ID 5'-most EST

1298_2.R1040 asn701132523.h1

Method BLASTN
NCBI GI g3986749
BLAST score 382
E value 0.0e+00
Match length 977
% identity 86

NCBI Description Hevea brasiliensis serine/threonine protein phosphatase

type 2A (PP2A) mRNA, complete cds

Seq. No.

1961

Contig ID

1298_3.R1040

5'-most EST

LIB3170-083-Q1-J1-F9

Method BLASTX
NCBI GI g3859548
BLAST score 368
E value 4.0e-35
Match length 66
% identity 100

NCBI Description (AF097182) protein phosphatase 2A catalytic subunit [Oryza

sativa]

Seq. No.

1962

Contig ID

1298 4.R1040

5'-most EST

LIB3170-083-Q1-K1-F10

Method BLASTN
NCBI GI g3986749
BLAST score 119
E value 4.0e-60
Match length 270
% identity 87

NCBI Description Hevea brasiliensis serine/threonine protein phosphatase

type 2A (PP2A) mRNA, complete cds

Seq. No.

1963

Contig ID

1298 5.R1040



```
jC-gmle01810089f02a1
5'-most EST
                  BLASTN
Method
                  g19600
NCBI GI
BLAST score
                   174
E value
                   9.0e-93
                   470
Match length
                   85
% identity
NCBI Description M.sativa mRNA for eIF-4D
                   1964
Seq. No.
                   1298 6.R1040
Contiq ID
                   LIB3049-054-Q1-E1-D11
5'-most EST
                   BLASTN
Method
                   g19701
NCBI GI
                   149
BLAST score
E value
                   4.0e-78
Match length
                   401
% identity
                   84
                   N.plumbaginifolia mRNA NeIF-5A2 for initiation factor 5A(2)
NCBI Description
                   1965
Seq. No.
                   1298 7.R1040
Contig ID
                   LIB3051-107-Q1-K1-C3
5'-most EST
                   BLASTN
Method
                   g3986749
NCBI GI
BLAST score
                   96
                   2.0e-46
E value
                   243
Match length
% identity
                   86
                   Hevea brasiliensis serine/threonine protein phosphatase
NCBI Description
                   type 2A (PP2A) mRNA, complete cds
                   1966
Seq. No.
                   1298_11.R1040
Contig ID
                   jex700904182.h1
 5'-most EST
                   BLASTN
Method
                   g3986749
NCBI GI
                   34
BLAST score
                   1.0e-09
 E value
                   66
 Match length
                   88
 % identity
                   Hevea brasiliensis serine/threonine protein phosphatase
 NCBI Description
                   type 2A (PP2A) mRNA, complete cds
                   1967
 Seq. No.
                   1298 12.R1040
 Contig ID
                   xpa700793188.h1
 5'-most EST
                   BLASTN
 Method
                    q2225884
 NCBI GI
                    123
 BLAST score
 E value
                    1.0e-62
                    339
 Match length
 % identity
                   Solanum tuberosum mRNA for eukaryotic initiation factor
 NCBI Description
```

Seq. No. 1968

5A5, complete cds



```
1299 1.R1040
Contig ID
                   djj700606009.h2
5'-most EST
                   BLASTX
Method
                   q2194139
NCBI GI
                   237
BLAST score
                   6.0e-20
E value
                   66
Match length
                   65
% identity
                   (AC002062) EST gb_ATTS0887 comes from this gene.
NCBI Description
                   [Arabidopsis thal\overline{i}ana]
                   1969
Seq. No.
                   1303 1.R1040
Contig ID
                   LIB3138-128-Q1-N1-B10
5'-most EST
Method
                   BLASTX
                   g2664210
NCBI GI
                   536
BLAST score
                   5.0e-55
E value
                   123
Match length
                   80
% identity
                   (AJ222644) asparaginyl-tRNA synthetase [Arabidopsis
NCBI Description
                   thaliana]
                   1970
Seq. No.
                   1303 2.R1040
Contig ID
                   LIB3<del>1</del>38-022-Q1-N1-F2
5'-most EST
                   BLASTN
Method
                   g210811
NCBI GI
                    127
BLAST score
                    1.0e-64
E value
                    634
Match length
                    88
% identity
                   Bean pod mottle virus coat protein gene, complete cds,
NCBI Description
                    complete middle component (M) RNA
                    1971
Seq. No.
                    1304 1.R1040
Contig ID
                    LIB3029-006-Q1-B1-G2
5'-most EST
                    BLASTX
Method
                    q691752
NCBI GI
                    1150
BLAST score
                    1.0e-126
E value
                    424
Match length
                    54
 % identity
                   (D29803) preproMP27-MP32 [Cucurbita sp.]
NCBI Description
                    1972
 Seq. No.
 Contig ID
                    1306 1.R1040
                    vwf7\overline{0}0676455.h1
 5'-most EST
                    BLASTX
 Method
 NCBI GI
                    q872116
                    1674
 BLAST score
                    0.0e + 00
 E value
                    580
 Match length
```

(X79770) sti (stress inducible protein) [Glycine max]

61

% identity

NCBI Description



```
1973
Seq. No.
                  1306 2.R1040
Contig ID
                  txt700732045.h1
5'-most EST
Method
                  BLASTX
                  g4115918
NCBI GI
                  502
BLAST score
                  2.0e-50
E value
                  154
Match length
% identity
                   69
                   (AF118222) similar to nascent polypeptide associated
NCBI Description
                  complex alpha chain [Arabidopsis thaliana]
                   1974
Seq. No.
                   1306 3.R1040
Contig ID
5'-most EST
                   dpv701101850.h1
Method
                   BLASTN
                   q872115
NCBI GI
                   93
BLAST score
                   2.0e-44
E value
Match length
                   277
                   83
% identity
                   G.max gmsti mRNA
NCBI Description
                   1975
Seq. No.
                   1306 4.R1040
Contig ID
                   jC-gmro02910039g01d1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4115918
BLAST score
                   207
                   2.0e-16
E value
                   70
Match length
                   63
% identity
                   (AF118222) similar to nascent polypeptide associated
NCBI Description
                   complex alpha chain [Arabidopsis thaliana]
                   1976
Seq. No.
                   1306 5.R1040
Contig ID
                   asj700967426.hl
5'-most EST
                   BLASTX
Method
                   q4115918
NCBI GI
BLAST score
                   325
                   4.0e-30
E value
                   77
Match length
                   82
% identity
                   (AF118222) similar to nascent polypeptide associated
NCBI Description
                   complex alpha chain [Arabidopsis thaliana]
                   1977
Seq. No.
                   1307 1.R1040
Contig ID
 5'-most EST
                   awf700840576.h1
Method
                   BLASTX
                   q2275196
NCBI GI
```

Method BLASTX
NCBI GI g2275196
BLAST score 475
E value 8.0e-49
Match length 132
% identity 81

NCBI Description (AC002337) water stress-induced protein, WSI76 isolog



```
[Arabidopsis thaliana]
                  1978
Seq. No.
                  1309 1.R1040
Contig ID
                  uC-gmflminsoy045g04b1
5'-most EST
Method
                  BLASTX
                   q3063713
NCBI GI
                   193
BLAST score
                   2.0e-14
E value
                   122
Match length
% identity
                  (AL022537) putative protein [Arabidopsis thaliana]
NCBI Description
                   1979
Seq. No.
                   1313 1.R1040
Contig ID
                   uC-gmrominsoy060b07b1
5'-most EST
                   1980
Seq. No.
                   1313 2.R1040
Contig ID
                   jC-gmf102220108a04d1
5'-most EST
                   1981
Seq. No.
                   1313 3.R1040
Contig ID
                   LIB3106-006-Q1-K1-B4
5'-most EST
                   BLASTN
Method
                   q4038468
NCBI GI
BLAST score
                   157
                   1.0e-82
E value
                   409
Match length
% identity
                   86
NCBI Description Porteresia coarctata histone H3 mRNA, complete cds
                   1982
Seq. No.
                   1313 4.R1040
Contig ID
                   jC-gmf102220052g08d1
5'-most EST
                   BLASTX
Method
                   g1200205
NCBI GI
                   725
BLAST score
                   3.0e-76
E value
                   217
Match length
                   65
 % identity
                   (X95753) DAG [Antirrhinum majus]
NCBI Description
                   1983
 Seq. No.
                   1313 5.R1040
```

Contig ID

 $jC-g\overline{m}le01810001g08d1$ 5'-most EST

BLASTN Method g4038468 NCBI GI 161 BLAST score 6.0e-85 E value Match length 409 % identity 86

NCBI Description Porteresia coarctata histone H3 mRNA, complete cds

1984 Seq. No.

1314 1.R1040 Contig ID 5'-most EST crh700852056.hl



Method BLASTN
NCBI GI g170023
BLAST score 112
E value 6.0e-56
Match length 112
% identity 100

NCBI Description Glycine max maturation-associated protein (MAT9) mRNA, complete cds

Seq. No. 1985

Contig ID 1318 1.R1040

5'-most EST LIB3170-046-Q1-J1-E6

Method BLASTX
NCBI GI g2501578
BLAST score 1322
E value 1.0e-146
Match length 308
% identity 85

NCBI Description ETHYLENE-INDUCIBLE PROTEIN HEVER >gi 2129913 pir S60047

ethylene-responsive protein 1 - Para rubber tree

>gi 1209317 (M88254) ethylene-inducible protein [Hevea

brasiliensis]

Seq. No. 1986

Contig ID 1318_3.R1040 5'-most EST k11701209129.h1

Method BLASTX
NCBI GI g2501578
BLAST score 409
E value 4.0e-40
Match length 88
% identity 94

NCBI Description ETHYLENE-INDUCIBLE PROTEIN HEVER >gi_2129913_pir__S60047

ethylene-responsive protein 1 - Para rubber tree

>gi_1209317 (M88254) ethylene-inducible protein [Hevea

brasiliensis]

Seq. No. 1987

Contig ID 1319_1.R1040 5'-most EST kl1701214022.h1

Seq. No. 1988

Contig ID 1319 2.R1040

5'-most EST uC-gmflminsoy119b03b1

Seq. No. 1989

Contig ID 1321 1.R1040

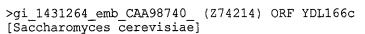
5'-most EST LIB3051-016-Q1-E1-H10

Method BLASTX
NCBI GI g2131352
BLAST score 289
E value 1.0e-25
Match length 136
% identity 39

NCBI Description hypothetical protein YDL166c - yeast (Saccharomyces cerevisiae) >gi 1061273 emb CAA91580 (Z67750) putative

protein [Saccharomyces cerevisiae]





1990 Seq. No.

1322 1.R1040 Contig ID

5'-most EST LIB3027-007-Q1-B1-A5

Seq. No.

Contig ID

1323 1.R1040

5'-most EST

LIB3106-098-Q1-K1-A6

Seq. No.

1992

Contig ID 5'-most EST 1331 1.R1040 zpv700759728.h1

Method NCBI GI BLASTX q2129842

BLAST score

157

E value

5.0e-10

Match length % identity

58 50

NCBI Description

SE60 protein - soybean >gi 509769 emb CAA79164 (Z18359) seed-specific low molecular weight sulfur-rich protein

[Glycine max]

Seq. No.

1993

Contig ID 5'-most EST 1331 2.R1040 ssr700560357.hl

Method

BLASTX

NCBI GI

q1667373

BLAST score E value

1785 0.0e+00

Match length

371

% identity

NCBI Description

(X82270) protein kinase [Medicago sativa]

Seq. No.

1994

Contig ID

1331 3.R1040

5'-most EST

LIB3051-054-Q1-K2-F4

Seq. No.

1995

Contig ID

1331 5.R1040

5'-most EST

LIB3072-027-Q1-E1-C8

Method NCBI GI BLASTX q2129842

BLAST score

E value

157

Match length

3.0e-10

% identity

58 50

NCBI Description

SE60 protein - soybean >gi 509769 emb CAA79164 (Z18359)

seed-specific low molecular weight sulfur-rich protein

[Glycine max]

Seq. No.

1996

Contig ID

1331 7.R1040

5'-most EST

LIB3051-031-Q1-K1-C10

Method NCBI GI BLASTN g1667372



105 BLAST score 1.0e-51 E value 240 Match length 86 % identity

NCBI Description M.sativa MMK4 mRNA for protein kinase

1997 Seq. No.

1331 8.R1040 Contig ID

LIB3170-061-Q1-J1-F1 5'-most EST

1998 Seq. No.

1331 9.R1040 Contig ID

LIB3051-031-Q1-K1-C11 5'-most EST

BLASTN Method g1667372 NCBI GI BLAST score 75 8.0e-34 E value Match length 91

96 % identity

NCBI Description M.sativa MMK4 mRNA for protein kinase

Seq. No. 1999

1331 11.R1040 Contig ID

jC-gmro02910024c08a1 5'-most EST

BLASTN Method g1667372 NCBI GI BLAST score 133 2.0e-68 E value Match length 293 86 % identity

NCBI Description M.sativa MMK4 mRNA for protein kinase

Seq. No.

1331 12.R1040 Contig ID

LIB3051-064-Q1-K1-F1 5'-most EST

2000

2001 Seq. No.

1331 14.R1040 Contig ID

 $jC-g\overline{m}st02400060f10d1$ 5'-most EST

2002 Seq. No.

1331 15.R1040 Contig ID uC-gmropic063f01b1 5'-most EST

BLASTN Method g1667372 NCBI GI 95 BLAST score 7.0e-46 E value 250 Match length 83 % identity

NCBI Description M.sativa MMK4 mRNA for protein kinase

2003 Seq. No.

1331 17.R1040 Contig ID

LIB3170-051-Q1-J1-B10 5'-most EST

Seq. No. 2004

1331 22.R1040 Contig ID

100 100 7.00



```
fua701042265.hl
5'-most EST
                  2005
Seq. No.
                  1331 23.R1040
Contig ID
                  LIB3051-012-Q1-E1-C2
5'-most EST
                  2006
Seq. No.
                  1331 24.R1040
Contig ID
                  LIB3051-054-Q1-K2-B8
5'-most EST
                  2007
Seq. No.
                  1333 1.R1040
Contig ID
                  LIB3039-018-Q1-E1-G5
5'-most EST
                  BLASTX
Method
                   g3860259
NCBI GI
BLAST score
                   189
                   3.0e-14
E value
                   105
Match length
                   42
% identity
                  (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
                   2008
Seq. No.
                   1338 1.R1040
Contig ID
                   LIB3109-034-Q1-K1-C2
5'-most EST
                   BLASTN
Method
                   g3025860
NCBI GI
                   35
BLAST score
                   7.0e-10
E value
                   111
Match length
                   92
% identity
NCBI Description Medicago truncatula Mt4 genomic sequence
                   2009
Seq. No.
                   1341_1.R1040
Contig ID
                   LIB3167-078-P1-K2-B7
5'-most EST
                   BLASTX
Method
                   q3914472
NCBI GI
                   526
BLAST score
                   2.0e-53
E value
                   129
Match length
                   78
 % identity
                   PHOTOSYSTEM II 10 KD POLYPEPTIDE PRECURSOR (PII10)
NCBI Description
                   >gi_322764_pir__S32021 photosystem II 10K protein - common
                   tobacco >gi_22669_emb_CAA49693_ (X70088) NtpII10 [Nicotiana
                   tabacum]
                   2010
 Seq. No.
                   1341 2.R1040
 Contig ID
                   fde700876719.h1
 5'-most EST
                   BLASTX
 Method
 NCBI GI
                   g3914472
                    318
 BLAST score
                    2.0e-29
 E value
 Match length
                   99
                    67
 % identity
                   PHOTOSYSTEM II 10 KD POLYPEPTIDE PRECURSOR (PII10)
 NCBI Description
```

>gi_322764_pir__S32021 photosystem II 10K protein - common



tobacco >gi_22669_emb_CAA49693_ (X70088) NtpII10 [Nicotiana tabacum]

Seq. No. 2011

Contig ID 1341_3.R1040

5'-most EST LIB3093-012-Q1-K1-A8

Method BLASTX
NCBI GI g3914472
BLAST score 272
E value 6.0e-24
Match length 126
% identity 41

NCBI Description PHOTOSYSTEM II 10 KD POLYPEPTIDE PRECURSOR (PII10)

>gi_322764_pir__S32021 photosystem II 10K protein - common tobacco >gi_22669_emb_CAA49693_ (X70088) NtpII10 [Nicotiana

tabacum]

Seq. No. 2012

Contig ID 1341_5.R1040

5'-most EST uC-gmropic043e05b1

Method BLASTN
NCBI GI g1418983
BLAST score 59
E value 2.0e-24
Match length 167
% identity 84

NCBI Description L.esculentum mRNA for photosystem II 10 kD protein

Seq. No. 2013

Contig ID 1341_6.R1040

5'-most EST LIB3040-027-Q1-E1-A2

Method BLASTX
NCBI GI g3914472
BLAST score 391
E value 1.0e-37
Match length 89
% identity 83

NCBI Description PHOTOSYSTEM II 10 KD POLYPEPTIDE PRECURSOR (PII10)

>gi_322764_pir__S32021 photosystem II 10K protein - common tobacco >gi_22669_emb_CAA49693_ (X70088) NtpII10 [Nicotiana

tabacum]

Seq. No.

Contig ID 1343_1.R1040

5'-most EST LIB3139-052-P1-N1-F3

2014

Method BLASTX
NCBI GI g2982303
BLAST score 1215
E value 1.0e-133
Match length 267
% identity 84

NCBI Description (AF051236) hypothetical protein [Picea mariana]

Seq. No. 2015

Contig ID 1343 2.R1040

5'-most EST jC-gmle01810024e10d1

Method BLASTX

% identity

NCBI Description

43

[Spinacia oleracea]



```
q4204315
NCBI GI
                  801
BLAST score
                  0.0e + 00
E value
Match length
                  717
% identity
                   50
                   (AC003027) Unknown protein [Arabidopsis thaliana]
NCBI Description
                  2016
Seq. No.
                   1343 4.R1040
Contig ID
                  bth700846788.h1
5'-most EST
                   BLASTX
Method
                   g4128133
NCBI GI
                   377
BLAST score
                   5.0e-36
E value
Match length
                   131
% identity
                   55
                  (AJ006068) dTDP-D-glucose 4,6-dehydratase [Homo sapiens]
NCBI Description
                   2017
Seq. No.
                   1343 5.R1040
Contiq ID
                   LIB3106-028-Q1-K1-H1
5'-most EST
                   BLASTX
Method
                   g3337435
NCBI GI
                   158
BLAST score
                   3.0e-23
E value
                   172
Match length
                   32
% identity
                   (AF060198) PsbY precursor; putative photosytem II peptide
NCBI Description
                   [Spinacia oleracea]
                   2018
Seq. No.
                   1343 7.R1040
Contig ID
                   LIB3073-016-Q1-K1-C7
5'-most EST
                   BLASTX
Method
                   q3337435
NCBI GI
BLAST score
                   265
E value
                   4.0e-23
                   131
Match length
                   38
% identity
                   (AF060198) PsbY precursor; putative photosytem II peptide
NCBI Description
                   [Spinacia oleracea]
                   2019
Seq. No.
                   1343 11.R1040
Contig ID
                   vzy700753763.h1
5'-most EST
                   2020
Seq. No.
                   1343 12.R1040
Contig ID
                   jC-qmle01810075b12a1
 5'-most EST
Method
                   BLASTX
                   q3337435
NCBI GI
                   189
BLAST score
                   3.0e-14
E value
                   56
Match length
```

448

(AF060198) PsbY precursor; putative photosytem II peptide



2021 Seq. No.

1343 14.R1040 Contig ID taw700656383.h1 5'-most EST

2022 Seq. No.

1343 15.R1040 Contig ID zhf700963687.h1 5'-most EST

2023 Seq. No.

1343 16.R1040 Contig ID hrw701057276.hl 5'-most EST

2024 Seq. No.

Contig ID 1344 1.R1040

LIB3051-026-Q1-K1-B3 5'-most EST

BLASTX Method g3747050 NCBI GI 411 BLAST score 5.0e-40 E value Match length 101 80 % identity

(AF093540) ribosomal protein L26 [Zea mays] NCBI Description

2025 Seq. No.

1344 2.R1040 Contig ID

LIB3106-093-Q1-K1-H2 5'-most EST

BLASTN Method g3747049 NCBI GI BLAST score 92 5.0e-44 E value 224 Match length 85

% identity

NCBI Description Zea mays ribosomal protein L26 mRNA, partial cds

2026 Seq. No.

1345 1.R1040 Contig ID kl1701215182.h1 5'-most EST

BLASTX Method NCBI GI q4337178 707 BLAST score 1.0e-74 E value 155 Match length 85 % identity

(AC006416) T31J12.5 [Arabidopsis thaliana] NCBI Description

2027 Seq. No.

1345 2.R1040 Contig ID

uC-gmrominsoy265e12b1 5'-most EST

Method BLASTX q4337178 NCBI GI 348 BLAST score 6.0e-33 E value 76 Match length 88 % identity

(AC006416) T31J12.5 [Arabidopsis thaliana] NCBI Description



```
2028
Seq. No.
                  1348 1.R1040
Contig ID
                  LIB3027-006-Q1-B1-A10
5'-most EST
                  BLASTN
Method
NCBI GI
                   g3241920
BLAST score
                   41
                   1.0e-13
E value
                   81
Match length
                   88
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MAE1, complete sequence [Arabidopsis thaliana]
                   2029
Seq. No.
                   1350 1.R1040
Contig ID
                   gsv701052806.h1
5'-most EST
Method
                   BLASTX
                   g4539307
NCBI GI
                   229
BLAST score
                   1.0e-18
E value
                   62
Match length
                   71
% identity
                   (AL049480) putative acidic ribosomal protein [Arabidopsis
NCBI Description
                   thaliana]
                   2030
Seq. No.
                   1350 2.R1040
Contig ID
                   g4397512
5'-most EST
                   BLASTX
Method
                   q4539307
NCBI GI
                   260
BLAST score
                   2.0e-22
E value
                   68
Match length
                   75
% identity
                   (AL049480) putative acidic ribosomal protein [Arabidopsis
NCBI Description
                   thaliana]
                   2031
Seq. No.
                   1352_1.R1040
Contig ID
                   jsh7\overline{0}1063775.h1
5'-most EST
                   BLASTN
Method
                   q555973
NCBI GI
                    269
BLAST score
                   1.0e-149
E value
                    692
Match length
 % identity
NCBI Description Pisum sativum 14-3-3-like protein mRNA, complete cds
                    2032
 Seq. No.
                    1352 2.R1040
 Contig ID
                    LIB3093-015-Q1-K1-C6
 5'-most EST
 Method
                    BLASTN
                    q555973
 NCBI GI
                    313
 BLAST score
                    1.0e-175
 E value
 Match length
                    786
 % identity
                    86
```

NCBI Description Pisum sativum 14-3-3-like protein mRNA, complete cds



```
2033
Seq. No.
                  1352 3.R1040
Contig ID
                  LIB3167-004-P1-K1-A2
5'-most EST
Method
                  BLASTN
                  g555973
NCBI GI
BLAST score
                  224
                  1.0e-122
E value
                  706
Match length
                  86
% identity
NCBI Description Pisum sativum 14-3-3-like protein mRNA, complete cds
                   2034
Seq. No.
                   1352 4.R1040
Contig ID
5'-most EST
                   zhf700963441.h1
Method
                   BLASTN
                   g2921511
NCBI GI
BLAST score
                   70
                   4.0e-31
E value
                   146
Match length
                   87
% identity
                  Fritillaria agrestis GF14 protein (GRF) mRNA, complete cds
NCBI Description
                   2035
Seq. No.
                   1352 5.R1040
Contig ID
5'-most EST
                   LIB3055-011-Q1-N1-A8
                   BLASTX
Method
                   q4544412
NCBI GI
                   303
BLAST score
                   2.0e-27
E value
                   82
Match length
                   70
% identity
                   (AC006955) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   2036
Seq. No.
                   1352 6.R1040
Contig ID
                   LIB3139-084-P1-N1-A1
5'-most EST
                   BLASTN
Method
                   q555973
NCBI GI
BLAST score
                   56
                   1.0e-22
E value
                   353
Match length
                   86
% identity
NCBI Description Pisum sativum 14-3-3-like protein mRNA, complete cds
                   2037
Seq. No.
                   1352 9.R1040
Contig ID
                   q4396272
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4544412
                   464
BLAST score
                   2.0e-46
E value
Match length
                   111
```

% identity NCBI Description

ion (AC006955) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2038

75

% identity

91 NCBI Description L.polyphyllus pPLB07 mRNA



```
1352 13.R1040
Contig ID
                  LIB3039-012-Q1-E1-A6
5'-most EST
                  BLASTN
Method
NCBI GI
                  g2921511
BLAST score
                  69
                  1.0e-30
E value
                  69
Match length
                  100
% identity
NCBI Description Fritillaria agrestis GF14 protein (GRF) mRNA, complete cds
                  2039
Seq. No.
                  1353 1.R1040
Contig ID
                  kmv700741661.h1
5'-most EST
                  BLASTN
Method
NCBI GI
                  g2264310
BLAST score
                  48
                  1.0e-17
E value
                  279
Match length
                  86
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MKP11, complete sequence [Arabidopsis thaliana]
                  2040
Seq. No.
                  1355 1.R1040
Contig ID
                  LIB3109-017-Q1-K1-D12
5'-most EST
Method
                  BLASTN
                  g2267138
NCBI GI
                  177
BLAST score
                   2.0e-94
E value
                   365
Match length
                   87
% identity
NCBI Description Prunus armeniaca ubiquitin-conjugating enzyme mRNA, partial
                   2041
Seq. No.
                   1355 2.R1040
Contig ID
                   jex700908608.h1
5'-most EST
                   BLASTN
Method
                   q2267138
NCBI GI
BLAST score
                   145
E value
                   1.0e-75
                   293
Match length
                   87
% identity
                  Prunus armeniaca ubiquitin-conjugating enzyme mRNA, partial
NCBI Description
                   cds
                   2042
Seq. No.
                   1355 3.R1040
Contig ID
                   LIB3051-048-Q1-K1-F7
5'-most EST
Method
                   BLASTN
                   q19497
NCBI GI
                   105
BLAST score
                   9.0e-52
E value
                   207
Match length
```



```
2043
Seq. No.
                  1355 4.R1040
Contig ID
                  LIB3049-054-Q1-E1-B4
5'-most EST
Method
                  BLASTX
NCBI GI
                  q1174162
                  534
BLAST score
                   1.0e-54
E value
                   105
Match length
% identity
                   90
                   (U44976) ubiquitin-conjugating enzyme [Arabidopsis
NCBI Description
                   thaliana] >gi_3746915 (AF091106) E2
                   ubiquitin-conjugating-like enzyme [Arabidopsis thaliana]
                   2044
Seq. No.
                   1355_6.R1040
Contig ID
                   uC-gmrominsoy050f02b1
5'-most EST
                   2045
Seq. No.
                   1356 1.R1040
Contig ID
                   LIB3072-060-Q1-K1-G5
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2832620
BLAST score
                   562
                   8.0e-58
E value
Match length
                   175
% identity
                   61
                   (AL021711) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2046
                   1363 1.R1040
Contig ID
                   LIB3\overline{1}39-068-P1-N1-D3
5'-most EST
                   BLASTX
Method
                   q4432844
NCBI GI
                   394
BLAST score
                   7.0e-38
E value
                   207
Match length
% identity
                   41
                   (AC006283) unknown protein [Arabidopsis thaliana]
NCBI Description
                   2047
Seq. No.
                   1363 2.R1040
Contig ID
                   sat701010967.hl
5'-most EST
                   2048
Seq. No.
                   1364 1.R1040
Contig ID
                   fde700876671.hl
5'-most EST
                   BLASTX
Method
                   q4033838
NCBI GI
BLAST score
                   1220
                   1.0e-134
E value
Match length
                   410
% identity
                   61
                   (Y18550) sigma-like factor [Arabidopsis thaliana]
NCBI Description
                   2049
```

Seq. No.

1364 2.R1040 Contig ID

5'-most EST jC-gmf102220132ab11d1



```
BLASTX
Method
                  q4033838
NCBI GI
                  430
BLAST score
                  2.0e-42
E value
                  101
Match length
                  84
% identity
                  (Y18550) sigma-like factor [Arabidopsis thaliana]
NCBI Description
                  2050
Seq. No.
                   1364_3.R1040
Contig ID
                   jC-gmro02910008c07a1
5'-most EST
                   BLASTX
Method
                   g4033838
NCBI GI
                   317
BLAST score
                   4.0e-29
E value
                   185
Match length
                   45
% identity
                   (Y18550) sigma-like factor [Arabidopsis thaliana]
NCBI Description
                   2051
Seq. No.
                   1364 4.R1040
Contig ID
                   jC-gmle01810004d04a1
5'-most EST
                   BLASTX
Method
                   g4033838
NCBI GI
BLAST score
                   226
                   2.0e-18
E value
                   120
Match length
% identity
                   41
                  (Y18550) sigma-like factor [Arabidopsis thaliana]
NCBI Description
                   2052
Seq. No.
                   1367 1.R1040
Contig ID
                   LIB3027-005-Q1-B1-G11
5'-most EST
                   BLASTX
Method
                   g2225877
NCBI GI
BLAST score
                   211
E value
                   1.0e-16
                   76
Match length
                   53
% identity
                   (AB002406) TIP49 [Rattus norvegicus] >gi 4106528 (AF100694)
NCBI Description
                   Pontin52 [Mus musculus] >gi 4521276 dbj BAA76313.1_
                   (AB001581) DNA helicase p50 [Rattus norvegicus]
                   2053
Seq. No.
                   1368 1.R1040
Contig ID
5'-most EST
                   wvk700680956.hl
                   BLASTX
Method
                   g2191191
NCBI GI
                   513
BLAST score
E value
                   7.0e-52
Match length
                   166
                   61
% identity
                   (AF007271) A TM021B04.14 gene product [Arabidopsis
NCBI Description
                   thaliana]
                   2054
Seq. No.
```

1372 1.R1040

Contig ID



```
zsq701123846.h1
5'-most EST
                  BLASTX
Method
                  g3250679
NCBI GI
                   447
BLAST score
                   4.0e-44
E value
                   236
Match length
% identity
NCBI Description (AL024486) putative protein [Arabidopsis thaliana]
Seq. No.
                   1375 1.R1040
Contig ID
5'-most EST
                   hyd7\overline{0}0730420.h1
                   BLASTX
Method
                   q3334299
NCBI GI
                   1049
BLAST score
                   1.0e-114
E value
Match length
                   219
% identity
                   PROTEASOME, ALPHA SUBUNIT (MULTICATALYTIC ENDOPEPTIDASE
NCBI Description
                   COMPLEX ALPHA SUBUNIT) >gi_2315211_emb_CAA74725_ (Y14339)
                   proteasome alpha subunit [Lycopersicon esculentum]
                   2056
Seq. No.
                   1376 1.R1040
Contig ID
                   kl1701213854.hl
5'-most EST
                   BLASTX
Method
                   q3176874
NCBI GI
BLAST score
                   753
                   1.0e-132
E value
Match length
                   478
% identity
                   (AF065639) cucumisin-like serine protease [Arabidopsis
NCBI Description
                   thaliana]
                   2057
Seq. No.
                   1377 1.R1040
Contig ID
                   rca701002041.hl
5'-most EST
                   BLASTX
Method
                    g3193314
NCBI GI
                    906
 BLAST score
                    1.0e-97
E value
Match length
                    328
                    55
 % identity
                   (AF069299) contains similarity to Arabidopsis scarecrow
 NCBI Description
                    (GB:U62798) [Arabidopsis thaliana]
                    2058
 Seq. No.
                    1379 1.R1040
 Contig ID
                    awf700843166.h1
 5'-most EST
                    BLASTX
 Method
                    q2270990
 NCBI GI
                    200
 BLAST score
                    3.0e-15
 E value
```

116

44

NCBI Description (AF004807) dehydrin [Glycine max]

Match length

% identity



```
2059
Seq. No.
                   1385 1.R1040
Contig ID
```

5'-most EST LIB3050-002-Q1-E1-H4

Seq. No.

2060

Contig ID 1386 1.R1040 rca701000271.hl 5'-most EST

Method BLASTN q1498327 NCBI GI BLAST score 272 E value 1.0e-151 276 Match length 100 % identity

NCBI Description Glycine max actin (Soy70) gene, partial cds

Seq. No.

2061

Contig ID 5'-most EST 1386 2.R1040 trc700562963.h1

Method BLASTX q1531672 NCBI GI BLAST score 1602 E value 0.0e+00Match length 377 % identity 93

NCBI Description (U68461) actin [Striga asiatica]

Seq. No.

2062

Contig ID 5'-most EST 1386 3.R1040 hyd700730194.h1

Method BLASTX NCBI GI g4139264 BLAST score 1157 E value 0.0e+00Match length 376 96

% identity

NCBI Description (AF111812) actin [Brassica napus]

Seq. No.

2063

Contig ID 5'-most EST 1386 4.R1040 k11701206694.h1

Method BLASTX NCBI GI g1703108 BLAST score 1856 0.0e + 00E value Match length 369 % identity 97

NCBI Description

ACTIN 2/7 >gi 2129525 pir S71210 actin 2 - Arabidopsis thaliana >gi 2129528 pir S68107 actin 7 ~ Arabidopsis thaliana >gi 1049307 (U37281) actin-2 [Arabidopsis thaliana] >gi 1943863 (U27811) actin7 [Arabidopsis

thaliana]

Seq. No.

2064

Contig ID 1386 5.R1040

5'-most EST

LIB3094-054-Q1-K1-E3

Method BLASTN NCBI GI g1498339



BLAST score 498 E value 0.0e+00Match length 1018 95 % identity

NCBI Description Glycine max actin (Soy118) gene, partial cds

Seq. No. Contig ID 2065

1386 6.R1040

5'-most EST

uC-gmrominsoy035g05b1

Method BLASTN NCBI GI q1498333 BLAST score 540 E value 0.0e+00Match length 1020 % identity 96

NCBI Description Glycine max actin (Soy57) gene, partial cds

Seq. No.

2066

Contig ID 5'-most EST 1391 1.R1040 sat701008458.hl

Method BLASTX NCBI GI q3337356 BLAST score 2273 0.0e+00 E value Match length 476 93 % identity

(AC004481) putative protein transport protein SEC61 alpha NCBI Description

subunit [Arabidopsis thaliana]

Seq. No.

2067

Contig ID 5'-most EST 1391 2.R1040 $asn7\overline{0}1140282.h1$

Method BLASTX NCBI GI g2827544 BLAST score 243 E value 3.0e-20

57 Match length 81 % identity

(AL021635) HSP associated protein like [Arabidopsis NCBI Description

thaliana]

Seq. No.

2068

Contig ID

1391 3.R1040

5'-most EST

LIB3170-053-Q1-J1-B12

Method BLASTX NCBI GI q3337356 BLAST score 385 5.0e-37 E value Match length 76 97 % identity

NCBI Description (AC004481) putative protein transport protein SEC61 alpha

subunit [Arabidopsis thaliana]

Seq. No.

2069

Contig ID

1391 4.R1040

5'-most EST

jC-gmf102220063a06a1

Method

BLASTX



NCBI GI g3337356 BLAST score 628 E value 1.0e-65 Match length 126 % identity 98

NCBI Description (AC004481) putative protein transport protein SEC61 alpha

subunit [Arabidopsis thaliana]

Seq. No. 2070

Contig ID 1391 5.R1040

5'-most EST jC-gmro02910040c03a1 Method BLASTN

NCBI GI g3766106
BLAST score 47
E value 3.0e-17
Match length 135
% identity 84

NCBI Description Arabidopsis thaliana chromosome 1 BAC F9K20 sequence,

complete sequence [Arabidopsis thaliana]

Seq. No. 2071

Contig ID 1391 6.R1040

5'-most EST LIB3028-053-Q1-B1-G8

Method BLASTX
NCBI GI g3337356
BLAST score 350
E value 5.0e-33
Match length 70
% identity 97

NCBI Description (AC004481) putative protein transport protein SEC61 alpha

.

subunit [Arabidopsis thaliana]

Seq. No. 2072

Contig ID 1391 8.R1040

5'-most EST uC-gmrominsoy174e10b1

Seq. No. 2073

Contig ID 1391_9.R1040 5'-most EST hyd700725385.h1

Method BLASTX
NCBI GI g3337356
BLAST score 210
E value 8.0e-17
Match length 55
% identity 78

NCBI Description (AC004481) putative protein transport protein SEC61 alpha

subunit [Arabidopsis thaliana]

Seq. No. 2074

Contig ID 1391_10.R1040 5'-most EST sat701013338.h1

Method BLASTX
NCBI GI g2827544
BLAST score 213
E value 3.0e-17
Match length 50
% identity 82



NCBI Description (AL021635) HSP associated protein like [Arabidopsis thaliana]

Seq. No. 2075

Contig ID 1391_11.R1040 5'-most EST uaw700665063.h1

Method BLASTX
NCBI GI g2827544
BLAST score 241
E value 3.0e-20
Match length 54
% identity 87

NCBI Description (AL021635) HSP associated protein like [Arabidopsis

thaliana]

Seq. No. 2076

Contig ID 1391_18.R1040 5'-most EST dpv701098828.h1

Method BLASTX
NCBI GI g3834321
BLAST score 248
E value 4.0e-21
Match length 143
% identity 50

NCBI Description (AC005679) Strong similarity to F13P17.9 gi_3337356

transport protein SEC61 alpha subunit homolog from Arabidopsis thaliana BAC gb_AC004481. [Arabidopsis

thaliana]

Seq. No. 2077

Contig ID 1394_1.R1040

5'-most EST LIB3028-031-Q1-B1-C9

Method BLASTX
NCBI GI g3355468
BLAST score 536
E value 1.0e-54
Match length 123
% identity 89

NCBI Description (AC004218) putative ribosomal protein L35 [Arabidopsis

thaliana]

Seq. No. 2078

Contig ID 1394 2.R1040

5'-most EST LIB3093-022-Q1-K1-B1

Method BLASTX
NCBI GI g4545262
BLAST score 172
E value 5.0e-12
Match length 42
% identity 71

NCBI Description (AF118230) metallothionein-like protein [Gossypium

hirsutum]

Seq. No. 2079

Contig ID 1394 3.R1040

5'-most EST LIB3107-060-Q1-K1-C9

Method BLASTN



g4545261 NCBI GI 40 BLAST score 4.0e-13 E value Match length 142 % identity Gossypium hirsutum metallothionein-like protein mRNA, NCBI Description complete cds 2080 Seq. No. 1394 5.R1040 Contig ID LIB3093-052-Q1-K1-F5 5'-most EST BLASTX Method g3355468 NCBI GI 186 BLAST score E value 1.0e-13 Match length 50 % identity (AC004218) putative ribosomal protein L35 [Arabidopsis NCBI Description thaliana] 2081 Seq. No. 1397 1.R1040 Contig ID awf700836730.hl 5'-most EST BLASTX Method g126078 NCBI GI BLAST score 380 2.0e-73 E value Match length 248 % identity 51 LATE EMBRYOGENESIS ABUNDANT PROTEIN D-34 (LEA D-34) NCBI Description >gi_81554_pir__S04046 embryonic abundant protein gD-34 upland cotton >gi_18501_emb_CAA31594_ (X13206) D-34 Lea protein [Gossypium hirsutum] >gi_167385 (M19389) storage protein [Gossypium hirsutum] >gi_226556_prf__1601521F Lea D-34 gene [Saguinus oedipus] 2082 Seq. No. Contig ID 1400 1.R1040 LIB3051-088-Q1-K1-E7 5'-most EST BLASTX Method q2146740 NCBI GI BLAST score 435 5.0e-43 E value 133 Match length % identity inner mitochondrial membrane protein - Arabidopsis thaliana NCBI Description >gi 603056 (U18126) inner mitochondrial membrane protein

[Arabidopsis thaliana]

2083 Seq. No. 1401 1.R1040 Contig ID

uC-qmflminsoy016c08b1 5'-most EST

BLASTX Method g4490732 NCBI GI BLAST score 2599 0.0e+00E value Match length 666



```
% identity
                  (AL035709) phosphoenolpyruvate carboxykinase (ATP)-like
NCBI Description
                  protein [Arabidopsis thaliana]
                  2084
Seq. No.
Contig ID
                  1401 2.R1040
                  LIB3065-027-Q1-N1-G12
5'-most EST
                  BLASTX
Method
                  q2623310
NCBI GI
                  571
BLAST score
                  3.0e-58
E value
                  248
Match length
% identity
                   (AC002409) unknown protein [Arabidopsis thaliana]
NCBI Description
                  >gi 3402721 (AC004261) unknown protein [Arabidopsis
                  thaliana]
                   2085
Seq. No.
                   1401 3.R1040
Contig ID
                  LIB3051-018-Q1-E1-H5
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2623310
                   289
BLAST score
                   2.0e-25
E value
                   138
Match length
% identity
                   (AC002409) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 3402721 (AC004261) unknown protein [Arabidopsis
                   thaliana]
                   2086
Seq. No.
                   1401 4.R1040
Contig ID
                   uC-gmropic023d09b1
5'-most EST
                   BLASTX
Method
                   g2623310
NCBI GI
                   516
BLAST score
                   3.0e-52
E value
                   184
Match length
                   58
% identity
                   (AC002409) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 3402721 (AC004261) unknown protein [Arabidopsis
                   thaliana]
                   2087
Seq. No.
                   1401 5.R1040
Contig ID
                   LIB3106-087-Q1-K1-A9
 5'-most EST
                   BLASTX
Method
                   g2623310
NCBI GI
                   216
 BLAST score
                   3.0e-17
 E value
                   61
Match length
 % identity
                    (AC002409) unknown protein [Arabidopsis thaliana]
```

2088 Seq. No.

thaliana]

NCBI Description

>gi_3402721 (AC004261) unknown protein [Arabidopsis



Contig ID 1401_6.R1040

NCBI GI g2623310 BLAST score 295 E value 2.0e-26 Match length 86 % identity 66

NCBI Description (AC002409) unknown protein [Arabidopsis thaliana]

>gi_3402721 (AC004261) unknown protein [Arabidopsis

thaliana]

Seq. No. 2089

Contig ID 1401_7.R1040

5'-most EST LIB3170-075-Q1-K2-H11

Method BLASTX
NCBI GI g1172571
BLAST score 224
E value 2.0e-18
Match length 44

% identity 95

NCBI Description PHOSPHOENOLPYRUVATE CARBOXYKINASE (ATP)

>gi_1076277_pir__S52637 phosphoenolpyruvate carboxykinase

(ATP) (EC 4.1.1.49) - cucumber >gi_567102 (L31899) phosphoenolpyruvate carboxykinase [Cucumis sativus]

Seq. No. 2090

Contig ID 1401_8.R1040

5'-most EST jC-gmst02400036f12d2

Seq. No. 2091

Contig ID 1405 1.R1040

5'-most EST jC-gmst02400077a05a1

Method BLASTX
NCBI GI g3327957
BLAST score 229
E value 1.0e-18
Match length 94
% identity 46

NCBI Description (AF060490) TLS-associated protein TASR-2 [Mus musculus]

 $>gi_3327976$ (AF067730) TLS-associated protein TASR-2 [Homo

sapiens]

Seq. No. 2092

Contig ID 1405_2.R1040 5'-most EST eep700864535.h1

Seq. No. 2093

Contig ID 1405 3.R1040

5'-most EST LIB3027-004-Q1-B1-E1

Seq. No. 2094

Contig ID 1409 1.R1040 5'-most EST sat701013306.h1

Method BLASTX
NCBI GI g2980760
BLAST score 238



E value 2.0e-19 Match length 93

% identity 53

NCBI Description (AL022198) putative protein [Arabidopsis thaliana]

Seq. No. 2095

Contig ID 1409_2.R1040 5'-most EST zsg701119142.h1

Method BLASTX
NCBI GI g2980760
BLAST score 226
E value 2.0e-18
Match length 58
% identity 67

NCBI Description (AL022198) putative protein [Arabidopsis thaliana]

Seq. No. 2096

Contig ID 1409_3.R1040 5'-most EST pcp700995525.h1

Method BLASTX
NCBI GI g421855
BLAST score 1044
E value 1.0e-114
Match length 342
% identity 61

NCBI Description alanine--tRNA ligase (EC 6.1.1.7) - Arabidopsis thaliana

(fragment)

Seq. No. 2097

Contig ID 1409_5.R1040

5'-most EST LIB3092-041-Q1-K1-H8

Method BLASTX
NCBI GI g1673366
BLAST score 201
E value 2.0e-15
Match length 86
% identity 51

NCBI Description (Z22673) cytosolic tRNA-Ala synthetase [Arabidopsis

thaliana]

Seq. No. 2098

Contig ID 1411_1.R1040 5'-most EST zsg701129392.h1

Method BLASTN
NCBI GI g1372965
BLAST score 387
E value 0.0e+00
Match length 832
% identity 88

NCBI Description Vicia faba CREB-like protein mRNA, complete cds

Seq. No. 2099

Contig ID 1411_2.R1040

5'-most EST uC-gmronoir037g03b1

Method BLASTN NCBI GI g1372965

BLAST score 50



```
4.0e-19
E value
                  113
Match length
% identity
                  91
NCBI Description Vicia faba CREB-like protein mRNA, complete cds
                  2100
Seq. No.
                  1411 3.R1040
Contig ID
                  uaw700664458.hl
5'-most EST
                  BLASTN
Method
                  g1372965
NCBI GI
BLAST score
                  32
                  1.0e-08
E value
Match length
                  44
                  93
% identity
NCBI Description Vicia faba CREB-like protein mRNA, complete cds
                  2101
Seq. No.
Contig ID
                  1412 1.R1040
                  LIB3028-032-Q1-B1-A11
5'-most EST
                  BLASTX
Method
                   q2257756
NCBI GI
BLAST score
                   300
                   1.0e-26
E value
Match length
                   286
                   33
% identity
                  (U82815) nucleolar histone deacetylase HD2-p39 [Zea mays]
NCBI Description
                   >gi 3650466 (AF026917) histone deacetylase HD2-p39 [Zea
                   mays]
                   2102
Seq. No.
                   1412 2.R1040
Contig ID
                   uC-gmrominsoy192h04b1
5'-most EST
                   2103
Seq. No.
                   1412 3.R1040
Contig ID
                   LIB3139-021-P1-N1-D5
5'-most EST
                   2104
Seq. No.
                   1416 1.R1040
Contig ID
                   LIB3170-017-Q1-K1-G9
5'-most EST
                   BLASTX
Method
                   q1350956
NCBI GI
                   534
BLAST score
E value
                   2.0e-54
                   116
Match length
                   91
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S20 (S22)
```

Seq. No.

2105

Contig ID 1416_2.R1040

5'-most EST LIB3040-026-Q1-E1-B2

Method BLASTX
NCBI GI g1350956
BLAST score 556
E value 5.0e-57
Match length 116
% identity 93

% identity

NCBI Description

82



```
NCBI Description 40S RIBOSOMAL PROTEIN S20 (S22)
                   2106
Seq. No.
                   1416 4.R1040
Contig ID
                   pxt700941510.h1
5'-most EST
                   BLASTX
Method
                   q1350956
NCBI GI
                   303
BLAST score
                   2.0e-49
E value
                   117
Match length
                   85
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S20 (S22)
                   2107
Seq. No.
                   1416 5.R1040
Contig ID
                   k117\overline{0}1215312.h1
5'-most EST
                   BLASTX
Method
                   g1350956
NCBI GI
                   274
BLAST score
                   3.0e-24
E value
                   74
Match length
                   70
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S20 (S22)
                   2108
Seq. No.
                   1416 6.R1040
Contig ID
                   taw700655032.h1
5'-most EST
                   BLASTX
Method
                   g1350956
NCBI GI
BLAST score
                   194
                   4.0e-15
E value
                   48
Match length
                   81
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S20 (S22)
                   2109
Seq. No.
                   1418 1.R1040
Contig ID
                   g431\overline{3}692
5'-most EST
                   BLASTN
Method
NCBI GI
                   q20901
                   285
BLAST score
                   1.0e-159
E value
                   730
Match length
                   85
 % identity
NCBI Description Pisum sativum mRNA for manganese superoxide dismutase
                    2110
Seq. No.
                    1418 2.R1040
 Contig ID
                   bth700849612.h1
 5'-most EST
                    BLASTX
Method
                    q3287977
 NCBI GI
                    556
 BLAST score
                    3.0e-57
 E value
                    127
 Match length
```

SUPEROXIDE DISMUTASE PRECURSOR (MN) >gi 945044 (U30841)

manganese superoxide dismutase precursor [Pisum sativum]



2111 Seq. No. 1422 1.R1040 Contig ID LIB3167-050-P1-K1-F9 5'-most EST BLASTX Method q3776572 NCBI GI 449 BLAST score 2.0e-44 E value 171 Match length 59 % identity (AC005388) ESTs gb_R65052, gb_AA712146, gb_H76533, NCBI Description gb_H76282, gb_AA650771, gb_H76287, gb_AA650887, gb_N37383, gb_Z29721 and gb_Z29722 come from this gene. [Arabidopsis thaliana] Seq. No. 2112 1424 1.R1040 Contig ID gsv701056551.h1 5'-most EST BLASTX Method q3193303 NCBI GI 342 BLAST score 1.0e-31 E value 99 Match length 64 % identity (AF069298) similar to several proteins containing a tandem NCBI Description repeat region such as Plasmodium falciparum GGM tandem repeat protein (GB:U27807); partial CDS [Arabidopsis thaliana] 2113 Seq. No. 1425 1.R1040 Contig ID LIB3107-053-Q1-K1-E3 5'-most EST BLASTX Method g3176690 NCBI GI BLAST score 1639 0.0e + 00E value 377 Match length % identity (AC003671) Similar to ubiquitin ligase gb_D63905 from S. NCBI Description cerevisiae. EST gb R65295 comes from this gene. [Arabidopsis thaliana] 2114 Seq. No. 1425 2.R1040 Contig ID vwf700678792.h1 5'-most EST BLASTX Method g3176690 NCBI GI

Method BLASTX
NCBI GI g3176690
BLAST score 731
E value 2.0e-77
Match length 306
% identity 54

NCBI Description (AC003671) Similar to ubiquitin ligase gb_D63905 from S.

cerevisiae. EST gb_R65295 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 2115

Contig ID 1427_1.R1040



```
LIB3072-056-Q1-K1-A9
5'-most EST
Seq. No.
                  2116
                  1430 1.R1040
Contig ID
                  ncj700985904.hl
5'-most EST
                  BLASTX
Method
                  q3928869
NCBI GI
                   219
BLAST score
                   5.0e-17
E value
                   306
Match length
                   27
```

% identity (AF093420) Hsp70 binding protein HspBP1 [Homo sapiens] NCBI Description

2117 Seq. No.

Contig ID 1430 2.R1040

LIB3027-001-Q1-B1-G7 5'-most EST

2118 Seq. No.

1430 4.R1040 Contig ID

jC-gmro02910070c02a1 5'-most EST

2119 Seq. No.

1431 1.R1040 Contig ID

LIB3109-014-Q1-K1-E7 5'-most EST

2120 Seq. No.

1433 1.R1040 Contig ID

LIB3027-003-Q1-B1-E4 5'-most EST

BLASTX Method g3243274 NCBI GI 148 BLAST score 2.0e-09 E value 61 Match length 51

% identity

NCBI Description (AF072134) TCP3 [Arabidopsis thaliana]

2121 Seq. No.

1433 2.R1040 Contig ID

LIB3049-028-Q1-E1-F10 5'-most EST

2122 Seq. No.

1434 1.R1040 Contig ID

5'-most EST LIB3040-048-Q1-E1-F9

BLASTX Method NCBI GI g2852449 BLAST score 927 1.0e-167 E value Match length 398 77 % identity

(D88207) protein kinase [Arabidopsis thaliana] >gi_2947061 NCBI Description (AC002521) putative protein kinase [Arabidopsis thaliana]

2123 Seq. No.

1434 2.R1040 Contig ID

LIB3051-104-Q1-K1-G3 5'-most EST

BLASTX Method g2852449 NCBI GI



```
BLAST score
                  887
                   1.0e-95
E value
                  246
Match length
                  73
% identity
                   (D88207) protein kinase [Arabidopsis thaliana] >gi 2947061
NCBI Description
                   (AC002521) putative protein kinase [Arabidopsis thaliana]
                   2124
Seq. No.
                   1434 3.R1040
Contig ID
                   epx701106557.h1
5'-most EST
Method
                   BLASTX
                   g2852449
NCBI GI
BLAST score
                   532
                   3.0e-54
                   142
```

E value Match length % identity 75

(D88207) protein kinase [Arabidopsis thaliana] >gi_2947061 NCBI Description (AC002521) putative protein kinase [Arabidopsis thaliana]

2125 Seq. No.

1434 4.R1040 Contig ID zsg701126118.h1 5'-most EST

BLASTX Method NCBI GI g2852447 BLAST score 178 7.0e-13 E value Match length 50 68 % identity

(D88206) protein kinase [Arabidopsis thaliana] NCBI Description

Seq. No.

2126

1437 1.R1040 Contig ID

LIB3065-027-Q1-N1-G9 5'-most EST

BLASTX Method g1169200 NCBI GI 195 BLAST score 6.0e-15 E value Match length 59 % identity 66

DNA-DAMAGE-REPAIR/TOLERATION PROTEIN DRT111 PRECURSOR NCBI Description >gi_421829_pir__S33706 DNA-damage resistance protein -Arabidopsis thaliana >gi_166694 (M98455) [Arabidopsis thaliana recombination and DNA-damage resistance protein

(DRT111) mRNA, complete cds.], gene product [Arabidopsis

thaliana]

2127 Seq. No.

1438 1.R1040 Contig ID zhf700963184.hl 5'-most EST

Seq. No. 2128

1438 2.R1040 Contig ID

LIB3107-054-Q1-K1-E1 5'-most EST

Seq. No. Contig ID

5'-most EST

2129 1438 3.R1040

pmv700891037.h1



```
BLASTX
Method
                  q2623299
NCBI GI
BLAST score
                   480
                  5.0e-48
E value
Match length
                  143
% identity
                   64
NCBI Description (AC002409) hypothetical protein [Arabidopsis thaliana]
                   2130
Seq. No.
                   1438 4.R1040
Contig ID
                                                              ~,~,
                  LIB3093-024-Q1-K1-D1
5'-most EST
                   2131
Seq. No.
                   1439 1.R1040
Contig ID
                   LIB3028-009-Q1-B1-A11
5'-most EST
Method
                   BLASTX
                   g3257577
NCBI GI
                   165
BLAST score
                   5.0e-11
E value
                   196
Match length
% identity
                   (AP000005) 211aa long hypothetical protein [Pyrococcus
NCBI Description
                   horikoshii]
                   2132
Seq. No.
                   1439 3.R1040
Contig ID
                   LIB3094-033-Q1-K1-B3
5'-most EST
Seq. No.
                   2133
                   1439 4.R1040
Contig ID
                   LIB3170-075-Q1-K1-F8
5'-most EST
                   2134
Seq. No.
                   1442 1.R1040
Contig ID
                   LIB3027-003-Q1-B1-H12
5'-most EST
                   2135
Seq. No.
                   1446 1.R1040
Contig ID
                   LIB3051-038-Q1-K1-E9
 5'-most EST
                   BLASTN
Method
                   q3985954
NCBI GI
                   34
BLAST score
                   3.0e-09
E value
                   85
Match length
 % identity
                   91
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
 NCBI Description
                   MRG21, complete sequence [Arabidopsis thaliana]
 Seq. No.
                   2136
                   1451 1.R1040
 Contig ID
                   LIB3027-003-Q1-B1-B8
 5'-most EST
 Method
                   BLASTX
                   g4539009
 NCBI GI
 BLAST score
                   639
                   1.0e-66
 E value
```

182

60

Match length

% identity

NCBI Description



```
(AL049481) putative protein [Arabidopsis thaliana]
NCBI Description
                  2137
Seq. No.
                  1454 1.R1040
Contiq ID
                  jC-gmro02800031a06a1
5'-most EST
                  BLASTX
Method
                  q3461820
NCBI GI
                  414
BLAST score
                  2.0e-40
E value
                  97
Match length
                  75
% identity
                  (AC004138) unknown protein [Arabidopsis thaliana]
NCBI Description
                  2138
Seq. No.
Contig ID
                  1456 1.R1040
                  LIB3051-080-Q1-K1-E3
5'-most EST
                  BLASTX
Method
                   q3861189
NCBI GI
                   285
BLAST score
                   5.0e-25
E value
                   118
Match length
                   52
% identity
                  (AJ235272) 50S RIBOSOMAL PROTEIN L14 (rplN) [Rickettsia
NCBI Description
                   prowazekii]
                   2139
Seq. No.
                   1456 2.R1040
Contig ID
                   LIB3027-003-Q1-B1-C8
5'-most EST
                   2140
Seq. No.
                   1456 3.R1040
Contig ID
                   jex700906734.h1
5'-most EST
                   2141
Seq. No.
                   1457 1.R1040
Contig ID
5'-most EST
                   uC-qmrominsoy095h04b1
Method
                   BLASTN
                   q1008880
NCBI GI
BLAST score
                   753
E value
                   0.0e + 00
Match length
                   1479
                   89
% identity
                   Phaseolus vulgaris eukaryotic initiation factor 5 (eIF-5)
NCBI Description
                   gene, complete cds
                   2142
Seq. No.
                   1457 2.R1040
Contig ID
                   leu701149702.h1
5'-most EST
                   BLASTN
Method
                   g1008880
NCBI GI
BLAST score
                   111
                   3.0e-55
E value
                   212
Match length
                   88
% identity
                   Phaseolus vulgaris eukaryotic initiation factor 5 (eIF-5)
```

gene, complete cds



```
2143
Seq. No.
Contig ID
                  1457 3.R1040
                  LIB3170-005-Q1-K1-F5
5'-most EST
Method
                  BLASTN
                  q1008880
NCBI GI
                  56
BLAST score
                  1.0e-22
E value
                  120
Match length
                  87
% identity
                  Phaseolus vulgaris eukaryotic initiation factor 5 (eIF-5)
NCBI Description
                  gene, complete cds
                  2144
Seq. No.
                  1457 4.R1040
Contig ID
                  uC-gmropic059d08b1
5'-most EST
Method
                  BLASTN
                  q1008880
NCBI GI
                   60
BLAST score
                   5.0e-25
E value
                  120
Match length
% identity
                   88
                  Phaseolus vulgaris eukaryotic initiation factor 5 (eIF-5)
NCBI Description
                   gene, complete cds
                   2145
Seq. No.
                   1457 5.R1040
Contig ID
5'-most EST
                   fde700870816.h1
Seq. No.
                   2146
                   1457 7.R1040
Contig ID
5'-most EST
                   jC-gmst02400073d03d1
                   BLASTN
Method
                   g1008880
NCBI GI
BLAST score
                   67
                   3.0e-29
E value
Match length
                   178
                   90
% identity
                  Phaseolus vulgaris eukaryotic initiation factor 5 (eIF-5)
NCBI Description
                   gene, complete cds
                   2147
Seq. No.
                   1458 1.R1040
Contig ID
5'-most EST
                   LIB3109-021-Q1-K1-A4
                   BLASTX
Method
NCBI GI
                   g1694621
                   1838
BLAST score
                   0.0e + 00
E value
                   461
Match length
% identity
```

NCBI Description

Seq. No.

Contig ID 1461_1.R1040 5'-most EST smc700744304.h1

2148

Method BLASTX
NCBI GI g2827552
BLAST score 136

(D70895) 3-ketoacyl-CoA thiolase [Cucurbita sp.]



```
2.0e-15
E value
Match length
                  60
                  73
% identity
                  (AL021635) predicted protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  2149
                  1464 1.R1040
Contig ID
5'-most EST
                  jC-gmro02910001d10d1
Seq. No.
                   2150
Contig ID
                   1466 1.R1040
                   LIB3074-031-Q1-K2-E8
5'-most EST
                   BLASTX
Method
                   q4102634
NCBI GI
                   255
BLAST score
                   1.0e-21
E value
Match length
                   48
% identity
NCBI Description (AF014396) Snakin-1 [Solanum tuberosum]
                   2151
Seq. No.
                   1466 2.R1040
Contig ID
                   sat701006340.h1
5'-most EST
                   BLASTX
Method
                   g4102634
NCBI GI
BLAST score
                   255
E value
                   3.0e-22
                   48
Match length
% identity
                   81
                  (AF014396) Snakin-1 [Solanum tuberosum]
NCBI Description
                   2152
Seq. No.
                   1469 1.R1040
Contig ID
                   awf700840524.h1
5'-most EST
                   2153
Seq. No.
                   1469 2.R1040
Contig ID
                   uC-gmrominsoy244g10b1
5'-most EST
Seq. No.
                   2154
                   1469 3.R1040
Contig ID
                   LIB3073-011-Q1-K1-G3
5'-most EST
                   2155
Seq. No.
                   1470 1.R1040
Contig ID
                   xpa700797086.hl
5'-most EST
                   BLASTX
Method
                   g3478637
NCBI GI
BLAST score
                   355
                   2.0e-33
E value
Match length
                   181
                   40
 % identity
                   (AC005546) R29425_1 [Homo sapiens]
 NCBI Description
```

Seq. No. 2156

Contig ID 1473_1.R1040

5'-most EST LIB3027-005-Q1-B1-F8



```
2157
Seq. No.
                  1477 1.R1040
Contig ID
5'-most EST
                  ncj700983761.hl
Method
                  BLASTX
                  q4454009
NCBI GI
                  210
BLAST score
                  2.0e-16
E value
                  102
Match length
% identity
                  43
NCBI Description (AL035396) putative protein [Arabidopsis thaliana]
                   2158
Seq. No.
                   1477 2.R1040
Contig ID
                   uC-gmrominsoy226a03b1
5'-most EST
Method
                   BLASTX
                   g2252829
NCBI GI
                   165
BLAST score
                   3.0e-11
E value
                   97
Match length
% identity
                   38
NCBI Description (AF013293) A_IG005I10.6 gene product [Arabidopsis thaliana]
                   2159
Seq. No.
                   1477 3.R1040
Contig ID
                   k117\overline{0}1214748.h1
5'-most EST
                   2160
Seq. No.
                   1477 4.R1040
Contig ID
                   LIB3051-017-Q1-E1-G5
5'-most EST
                   2161
Seq. No.
                   1479 1.R1040
Contig ID
                   jC-gmle01810041h06a1
5'-most EST
                   BLASTX
Method
                   g2459448
NCBI GI
                   692
BLAST score
                   1.0e-105
E value
                   311
Match length
                   59
% identity
                   (AC002332) putative cinnamoyl-CoA reductase [Arabidopsis
NCBI Description
                   thaliana]
                   2162
Seq. No.
                   1479 2.R1040
Contig ID
                   LIB3106-057-Q1-K1-B9
5'-most EST
                   BLASTX
Method
                   q2459446
NCBI GI
BLAST score
                   322
                   9.0e-30
E value
Match length
                   145
                   60
% identity
                   (AC002332) putative cinnamoyl-CoA reductase [Arabidopsis
```

2163 Seq. No.

NCBI Description

1480 1.R1040 Contig ID

thaliana]

NCBI Description



```
LIB3040-028-Q1-E1-H7
5'-most EST
                  BLASTX
Method
                  q4539408
NCBI GI
                  389
BLAST score
                  4.0e-37
E value
                  221
Match length
% identity
                  (ALO49524) putative alpha NAC [Arabidopsis thaliana]
NCBI Description
                  2164
Seq. No.
                  1480 2.R1040
Contig ID
                  LIB3049-054-Q1-E1-F10
5'-most EST
                  BLASTX
Method
                  q4539408
NCBI GI
                  287
BLAST score
                  1.0e-25
E value
                  84
Match length
                  69
% identity
NCBI Description (AL049524) putative alpha NAC [Arabidopsis thaliana]
Seq. No.
                  2165
                  1480 3.R1040
Contig ID
                  LIB3040-014-Q1-E1-H10
5'-most EST
                  2166
Seq. No.
                  1480 4.R1040
Contig ID
                  crh700853176.h1
5'-most EST
                  BLASTX
Method
                  g4115918
NCBI GI
                   200
BLAST score
                   2.0e-15
E value
                   93
Match length
                   49
% identity
                  (AF118222) similar to nascent polypeptide associated
NCBI Description
                   complex alpha chain [Arabidopsis thaliana]
                   2167
Seq. No.
                   1480 7.R1040
Contig ID
                   fua701039265.h1
5'-most EST
                   2168
Seq. No.
                   1480 8.R1040
Contig ID
                   kl1701212553.h1
5'-most EST
                   2169
Seq. No.
                   1481 1.R1040
Contig ID
                   txt700736308.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1085622
BLAST score
                   1427
E value
                   1.0e-175
Match length
                   378
                   77
% identity
                   alcohol dehydrogenase (EC 1.1.1.1) 1F - Phaseolus
```

dehydrogenase-1F [Phaseolus acutifolius]

acutifolius >gi_452769_emb_CAA80691_ (Z23170) alcohol



```
2170
Seq. No.
                   1482 1.R1040
Contig ID
                   zsq701120954.h1
5'-most EST
Seq. No.
                   2171
Contig ID
                   1482 2.R1040
                   LIB3049-040-Q1-E1-F3
5'-most EST
                   2172
Seq. No.
                   1483 1.R1040
Contig ID
5'-most EST
                   LIB3065-011-Q1-N1-F6
                   BLASTX
Method
                   g1903034
NCBI GI
                   292
BLAST score
                   3.0e-26
E value
Match length
                   110
% identity
                  (X94625) amp-binding protein [Brassica napus]
NCBI Description
Seq. No.
                   1488 1.R1040
Contig ID
5'-most EST
                   LIB3049-031-Q1-E1-E3
                   BLASTX
Method
                   g1039355
NCBI GI
                   491
BLAST score
                   2.0e-49
E value
Match length
                   151
                   59
 % identity
NCBI Description (X92179) alcohol dehydrogenase [Solanum tuberosum]
                   2174
Seq. No.
                   1489 1.R1040
 Contig ID
                   LIB3139-044-P1-N1-E3
 5'-most EST
                   BLASTX
Method
                   g3264767
 NCBI GI
                   543
 BLAST score
                    5.0e-55
 E value
                   279
 Match length
                    43
 % identity
                   (AF071893) AP2 domain containing protein [Prunus armeniaca]
 NCBI Description
                    2175
 Seq. No.
                    1489 2.R1040
 Contig ID
                    ssr700556206.hl
 5'-most EST
                    BLASTX
 Method
                    g3264767
 NCBI GI
                    660
 BLAST score
                    7.0e-69
 E value
                    272
 Match length
 % identity
                    51
                   (AF071893) AP2 domain containing protein [Prunus armeniaca]
 NCBI Description
                    2176
 Seq. No.
                    1489 3.R1040
 Contig ID
```

qsv701056072.h1

BLASTX

q3264767

5'-most EST Method

NCBI GI



160 BLAST score 3.0e-10 E value 73 Match length 48 % identity (AF071893) AP2 domain containing protein [Prunus armeniaca] NCBI Description 2177 Seq. No. 1489 4.R1040 Contig ID LIB3093-038-Q1-K1-E4 5'-most EST 2178 Seq. No. 1489 6.R1040 Contig ID jC-gmf102220068b09d1 5'-most EST Seq. No. 2179 Contig ID 1492 1.R1040 LIB3040-030-Q1-E1-A10 5'-most EST Method BLASTX q3367515 NCBI GI BLAST score 912 E value 1.0e-139 363 Match length 71 % identity (AC004392) Similar to NCBI Description glucose-6-phosphate/phosphate-translocator (GPT) gb_AF020814 from Pisum sativum. [Arabidopsis thaliana] 2180 Seq. No. 1493 1.R1040 Contig ID LIB3027-002-Q1-B1-B8 5'-most EST Method BLASTX g4467147 NCBI GI 199 BLAST score 2.0e-15 E value 110 Match length % identity 51 (AL035540) putative protein [Arabidopsis thaliana] NCBI Description Seq. No. 2181 1493 2.R1040 Contig ID sat701013373.hl 5'-most EST BLASTX Method q4467147 NCBI GI 307 BLAST score E value 5.0e-34 159 Match length % identity (AL035540) putative protein [Arabidopsis thaliana] NCBI Description Seq. No. 2182 1495 1.R1040 Contig ID 5'-most EST crh700850359.h1 BLASTX Method

NCBI GI g2281631
BLAST score 317
E value 1.0e-28
Match length 201

√ ೬.



Seq. No. 2183

Contig ID 1495_2.R1040 5'-most EST awf700842114.h1

Method BLASTX
NCBI GI g2281631
BLAST score 305
E value 3.0e-27
Match length 140
% identity 52

NCBI Description (AF003096) AP2 domain containing protein RAP2.3

[Arabidopsis thaliana]

Seq. No. 2184

Contig ID 1495 3.R1040

5'-most EST LIB3170-067-Q1-K1-B1

Method BLASTX
NCBI GI g2281631
BLAST score 304
E value 2.0e-27
Match length 83
% identity 73

NCBI Description (AF003096) AP2 domain containing protein RAP2.3

[Arabidopsis thaliana]

Seq. No. 2185

Contig ID 1497_1.R1040 5'-most EST smc700749770.h1

Seq. No. 2186

Contig ID 1498_1.R1040 5'-most EST asj700967459.h1

Seq. No. 2187

Contig ID 1499_1.R1040 5'-most EST uxk700669903.h1

Seq. No. 2188

Contig ID 1500_1.R1040

5'-most EST LIB3028-048-Q1-B1-A3

Method BLASTX
NCBI GI g2213538
BLAST score 191
E value 2.0e-14
Match length 155
% identity 43

NCBI Description (X98740) DNA-binding protein PD2 [Pisum sativum]

Seq. No. 2189

Contig ID 1501_1.R1040

5'-most EST LIB3072-028-Q1-E1-B4

Method BLASTX NCBI GI g1531756 BLAST score 178



```
9.0e-13
E value
Match length
                  70
% identity
                 (X82413) proline-rich-like protein [Asparagus officinalis]
NCBI Description
Seq. No.
                  2190
                  1502 1.R1040
Contig ID
                  kl1701204348.h2
5'-most EST
                  BLASTX
Method
                  q3128174
NCBI GI
BLAST score
                  251
                  2.0e-21
E value
                  93
Match length
                   59
% identity
NCBI Description (AC004521) hypothetical protein [Arabidopsis thaliana]
                   2191
Seq. No.
                   1503 1.R1040
Contig ID
                   fua701041667.h1
5'-most EST
                   BLASTX
Method
                   q3850581
NCBI GI
BLAST score
                   648
                   2.0e-70
E value
                   185
Match length
% identity
                   (AC005278) EST gb N96383 comes from this gene. [Arabidopsis
NCBI Description
                   thaliana]
                   2192
Seq. No.
                   1507 1.R1040
Contig ID
                   LIB3028-053-Q1-B1-E2
5'-most EST
                   2193
Seq. No.
                   1510 1.R1040
Contig ID
                   LIB3049-030-Q1-E1-D6
5'-most EST
                   BLASTX
Method
                   g266743
NCBI GI
BLAST score
                   2213
                   0.0e + 00
E value
                   477
Match length
% identity
                   87
                   PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) /
NCBI Description
                   DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE-PROTEIN GLYCOTRANSFERASE
                   (GLYCOSYLATION SITE-BINDING CHAIN) (GSBP)
                   >gi_68467_pir__ISAASS protein disulfide-isomerase (EC
                   5.3.4.1) precursor (clone L1) - alfalfa
                   >gi 19654 emb CAA77575 (Z11499) protein disulfide
                   isomerase [Medicago sativa]
                   2194
Seq. No.
                   1510 2.R1040
Contig ID
 5'-most EST
                   ncj7\overline{0}0980254.h1
                   BLASTN
Method
                   g166417
NCBI GI
 BLAST score
                   97
```

5.0e-47

141

E value Match length



% identity

NCBI Description Alfalfa putative endomembrane protein mRNA, complete cds

Seq. No.

1513 2.R1040 Contig ID

LIB3027-001-Q1-B1-D2 5'-most EST

BLASTX Method q119640 NCBI GI 206 BLAST score 2.0e-16 E value 90 Match length 51 % identity

1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG (PROTEIN NCBI Description

E8) >gi 82109 pir _S01642 ripening protein E8 - tomato >gi 19199_emb_CAA31789_ (X13437) E8 protein [Lycopersicon

esculentum]

2196 Seq. No.

1514 1.R1040 Contig ID pxt700943448.h1 5'-most EST

2197 Seq. No.

1517 1.R1040 Contig ID zhf700955891.h1 5'-most EST

BLASTX Method q3335374 NCBI GI BLAST score 342 5.0e-32 E value 157 Match length % identity 52

(AC003028) glutaredoxin-like protein [Arabidopsis thaliana] NCBI Description

2198 Seq. No.

1518 1.R1040 Contig ID

LIB3051-068-Q1-K1-H1 5'-most EST

BLASTX Method g3023817 NCBI GI 566 BLAST score 6.0e-58 E value 152 Match length 72 % identity

GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE, CHOLOROPLAST ISOFORM NCBI Description

PRECURSOR (G6PD) >gi_1480344 emb CAA67782 (X99405) glucose-6-phosphate dehydrogenase [Nicotiana tabacum]

2199 Seq. No.

1518 2.R1040 Contig ID

uC-gmrominsoy311f07b1 5'-most EST

BLASTX Method q3023817 NCBI GI BLAST score 762 3.0e-81 E value 171 Match length 87 % identity

GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE, CHOLOROPLAST ISOFORM NCBI Description

PRECURSOR (G6PD) >gi 1480344_emb_CAA67782_ (X99405) glucose-6-phosphate dehydrogenase [Nicotiana tabacum]



2200

Seq. No.

```
1519 1.R1040
Contig ID
                  uC-gmropic058g03b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q3915847
BLAST score
                  898
                  6.0e-97
E value
                  180
Match length
                  94
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S2 >gi_2335095 (AC002339) putative
                  40S ribosomal protein S2 [Arabidopsis thaliana]
                  2201
Seq. No.
                  1519 2.R1040
Contig ID
                  uC-gmronoir034a10b1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3915847
                  1063
BLAST score
E value
                  1.0e-116
                  219
Match length
                  92
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S2 >gi_2335095 (AC002339) putative
                   40S ribosomal protein S2 [Arabidopsis thaliana]
                   2202
Seq. No.
                   1519 3.R1040
Contig ID
                   uC-gmrominsoy315d11b1
5'-most EST
                  BLASTX
Method
                   q3915847
NCBI GI
                   474
BLAST score
E value
                   2.0e-47
                   103
Match length
                   89
% identity
NCBI Description 40S RIBOSOMAL PROTEIN S2 >gi_2335095 (AC002339) putative
                   40S ribosomal protein S2 [Arabidopsis thaliana]
                   2203
Seq. No.
                   1519 4.R1040
Contig ID
                   jC-gmst02400009e05d1
 5'-most EST
                   BLASTX
Method
                   g3915847
NCBI GI
                   184
BLAST score
                   2.0e-13
E value
                   53
Match length
                   66
 % identity
NCBI Description 40S RIBOSOMAL PROTEIN S2 >gi 2335095 (AC002339) putative
                   40S ribosomal protein S2 [Arabidopsis thaliana]
                   2204
 Seq. No.
 Contig ID
                   1519 5.R1040
                   LIB3106-062-Q1-K1-D6
 5'-most EST
                   BLASTX
 Method
                   q3915847
 NCBI GI
                   295
 BLAST score
 E value
                   1.0e-26
                   67
 Match length
```



```
% identity
                   40S RIBOSOMAL PROTEIN S2 >gi_2335095 (AC002339) putative
NCBI Description
                    40S ribosomal protein S2 [Arabidopsis thaliana]
Seq. No.
                   2205
                   1523 1.R1040
Contig ID
                   LIB3072-047-Q1-K1-C3
5'-most EST
Method
                   BLASTX
NCBI GI
                    q125722
BLAST score
                    427
E value
                    8.0e-42
Match length
                    201
% identity
                    KUNITZ-TYPE TRYPSIN INHIBITOR KTI1 PRECURSOR
NCBI Description
                    >gi_81814_pir__JQ1091 trypsin inhibitor KTi1 (Kunitz) -
                    soy\overline{b}ean > \overline{gi}_2 \overline{56}635 \underline{bbs}_1 \overline{15}028 (S45035) Kunitz trypsin
                    inhibitor KTil [soybeans, Peptide, 203 aa] [Glycine max]
Seq. No.
                    2206
                    1523 3.R1040
Contig ID
5'-most EST
                    crh700854192.h1
Seq. No.
                    2207
Contig ID
                    1524 1.R1040
                    rca700999295.h1
5'-most EST
Method
                    BLASTX
                    g1935019
NCBI GI
BLAST score
                    835
E value
                    2.0e-89
                    330
Match length
                    53
% identity
                   (Z93774) sucrose transport protein [Vicia faba]
NCBI Description
```

Seq. No. 2208

1526 1.R1040 Contig ID

jC-gmf102220078e07a1 5'-most EST

BLASTX Method NCBI GI g4263711 BLAST score 323 8.0e-30 E value Match length 119 55 % identity

(AC006223) putative CCR4-associated transcription factor NCBI Description

[Arabidopsis thaliana]

2209 Seq. No.

1528 1.R1040 Contig ID

uC-gmflminsoy081g03b1 5'-most EST

Method BLASTX NCBI GI q3096919 BLAST score 550 E value 3.0e-56 Match length 125 81 % identity

(AL023094) putative serine/threonine protein kinase NCBI Description

[Arabidopsis thaliana]

NCBI Description



```
2210
Seq. No.
                  1529 1.R1040
Contig ID
                  LIB3040-041-Q1-E1-G7
5'-most EST
                  BLASTX
Method
                  q100200
NCBI GI
BLAST score
                  1080
                  1.0e-118
E value
                  273
Match length
                  75
% identity
NCBI Description chlorophyll a/b-binding protein type I precursor - tomato
Seq. No.
                   1529 2.R1040
Contig ID
                   LIB3040-026-Q1-E1-E6
5'-most EST
                   BLASTX
Method
NCBI GI
                   q100200
BLAST score
                   555
E value
                   6.0e-57
                   144
Match length
% identity
                   73
NCBI Description chlorophyll a/b-binding protein type I precursor - tomato
                   2212
Seq. No.
                   1529 3.R1040
Contig ID
                   LIB3138-088-P1-N1-H4
5'-most EST
                   BLASTN
Method
NCBI GI
                   q4538949
BLAST score
                   47
                   2.0e-17
E value
Match length
                   67
                   93
% identity
                   Arabidopsis thaliana DNA chromosome 4, BAC clone F24G24
NCBI Description
                   (ESSA project)
                   2213
Seq. No.
                   1529 4.R1040
Contig ID
                   leu701157323.h1
5'-most EST
                   BLASTX
Method
                   g100200
NCBI GI
                   398
BLAST score
                   1.0e-38
E value
                   118
Match length
                   66
% identity
NCBI Description chlorophyll a/b-binding protein type I precursor - tomato
Seq. No.
                   2214
                   1529 9.R1040
Contig ID
                   LIB3\overline{1}38-120-Q1-N1-F12
5'-most EST
Method
                   BLASTX
                   q4538963
NCBI GI
BLAST score
                   475
E value
                   2.0e-47
                   145
Match length
                   73
% identity
```

[Arabidopsis thaliana]

(AL049488) chlorophyll a/b-binding protein-like



```
2215
Seq. No.
                   1529 10.R1040
Contig ID
                   jC-g\overline{m}st02400078b01d1
5'-most EST
                   BLASTX
Method
                   g4538963
NCBI GI
                   349
BLAST score
                   7.0e-33
E value
                   80
Match length
                   88
% identity
                   (AL049488) chlorophyll a/b-binding protein-like
NCBI Description
                   [Arabidopsis thaliana]
                   2216
Seq. No.
                   1529 11.R1040
Contig ID
5'-most EST
                   ssr700556044.hl
Method
                   BLASTN
                   g1644288
NCBI GI
                   60
BLAST score
                   5.0e-25
E value
                   120
Match length
% identity
                   88
NCBI Description B.juncea mRNA for chlorophyll a/b-binding protein
                   2217
Seq. No.
                   1529 12.R1040
Contig ID
                   jC-gmst02400023f02d1
5'-most EST
                   2218
Seq. No.
                   1529 13.R1040
Contig ID
                   LIB3106-093-Q1-K1-A11
5'-most EST
                   BLASTX
Method
                   g100200
NCBI GI
                   514
BLAST score
                   3.0e-52
E value
Match length
                   114
                   63
% identity
                   chlorophyll a/b-binding protein type I precursor - tomato
NCBI Description
                   2219
Seq. No.
                   1530 1.R1040
Contig ID
                   jC-qmle01810041e09a1
5'-most EST
                   BLASTX
Method
                   q2407800
NCBI GI
BLAST score
                   506
E value
                   4.0e-51
                   110
Match length
% identity
                   (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
NCBI Description
                   2220
Seq. No.
                   1538 1.R1040
Contig ID
                   LIB3073-019-Q1-K1-G11
 5'-most EST
Method
                   BLASTX
                   g3935169
 NCBI GI
BLAST score
                    238
```

8.0e-20

68

E value Match length

```
% identity
NCBI Description
                  (AC004557) F17L21.12 [Arabidopsis thaliana]
Seq. No.
                   2221
Contig ID
                   1538 2.R1040
5'-most EST
                   jC-gmle01810020q02d1
Method
                   BLASTX
NCBI GI
                   g2739382
BLAST score
                   357
E value
                   2.0e-33
Match length
                   205
% identity
                   47
NCBI Description
                  (AC002505) myosin heavy chain-like protein [Arabidopsis
                   thaliana]
Seq. No.
                   2222
Contig ID
                   1538 3.R1040
5'-most EST
                   bth700845402.h1
Method
                   BLASTN
NCBI GI
                   q3228389
BLAST score
                   \bar{4}2
E value
                   3.0e-14
Match length
                   102
% identity
                   43
NCBI Description Genomic sequence for Arabidopsis thaliana BAC F17L21,
                   complete sequence [Arabidopsis thaliana]
Seq. No.
                   2223
Contig ID
                  1541 1.R1040
5'-most EST
                  LIB3106-102-Q1-K1-D12
Method
                  BLASTX
                   g2677830
NCBI GI
BLAST score
                   772
E value
                   3.0e-82
Match length
                  166
% identity
                   90
                  (U93168) ribosomal protein L12 [Prunus armeniaca]
NCBI Description
Seq. No.
                  2224
                  1541 2.R1040
Contig ID
5'-most EST
                  jC-gmro02910008a06d1
Method
                  BLASTX
NCBI GI
                  g4371282
BLAST score
                  236
E value
                  9.0e-20
Match length
                  52
% identity
                  81
                  (AC006260) putative 60S ribosomal protein L12 [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  2225
Contig ID
                  1543 1.R1040
5'-most EST
                  LIB3027-003-Q1-B1-C7
```

Seq. No.

2226

Contig ID 1548 1.R1040

5'-most EST

LIB3139-004-P1-N1-C7



Method BLASTX
NCBI GI g729335
BLAST score 864
E value 8.0e-93
Match length 240
% identity 64
NCBI Description SUCCINAL

NCBI Description SUCCINATE DEHYDROGENASE (UBIQUINONE) IRON-SULFUR PROTEIN PRECURSOR (IP) >gi_508849 (L27705) succinate dehydrogenase

iron-protein subunit [Drosophila melanogaster]

Seq. No. 2227

Contig ID 1552 1.R1040

5'-most EST LIB3109-020-Q1-K1-F3

Method BLASTN
NCBI GI g2353172
BLAST score 56
E value 2.0e-22
Match length 200
% identity 82

NCBI Description Arabidopsis thaliana sigma factor 2 (SIG2) mRNA, nuclear

gene encoding chloroplast protein, complete cds

Seq. No. 2228

Contig ID 1554 1.R1040

5'-most EST LIB3051-017-Q1-E1-C9

Method BLASTX
NCBI GI g417494
BLAST score 237
E value 9.0e-20
Match length 106
% identity 49

NCBI Description 18 KD SEED MATURATION PROTEIN >gi_170008 (M80666)

maturation polypeptide [Glycine max]

>gi_444332_prf__1906381A 18kD late embryogenesis abundant

protein [Glycine max]

Seq. No. 2229

Contig ID 1562_1.R1040

5'-most EST jC-gmro02910060a09a1

Method BLASTN
NCBI GI 94234940
BLAST score 316
E value 1.0e-177
Match length 950
% identity 83

NCBI Description Populus tremula x Populus tremuloides cytosolic

phosphoglucomutase (pgm) mRNA, complete cds

Seq. No. 2230

Contig ID 1562 2.R1040 5'-most EST smc700747567.h1

Method BLASTX
NCBI GI g4234941
BLAST score 353
E value 2.0e-33
Match length 83
% identity 86



NCBI Description (AF097938) cytosolic phosphoglucomutase [Populus tremula x Populus tremuloides]

Seq. No. 2231

Contig ID 1562 3.R1040

5'-most EST LIB3139-039-P1-N1-B9

Method BLASTN
NCBI GI g4234940
BLAST score 168
E value 2.0e-89
Match length 491
% identity 84

NCBI Description Populus tremula x Populus tremuloides cytosolic

phosphoglucomutase (pgm) mRNA, complete cds

Seq. No. 2232

Contig ID 1563_1.R1040 5'-most EST sat701003565.h1

Method BLASTX
NCBI GI g1353352
BLAST score 1572
E value 1.0e-175
Match length 462
% identity 64

NCBI Description (U31975) alanine aminotransferase [Chlamydomonas

reinhardtii]

Seq. No. 2233

Contig ID 1563 2.R1040

5'-most EST fC-gmle7000743719a1

Method BLASTX
NCBI GI g1353352
BLAST score 401
E value 6.0e-39
Match length 104
% identity 72

NCBI Description (U31975) alanine aminotransferase [Chlamydomonas

reinhardtii]

Seq. No. 2234

Contig ID 1563_3.R1040

5'-most EST LIB3073-025-Q1-K1-B2

Seq. No. 2235

Contig ID 1563_4.R1040 5'-most EST trc700564493.h1

Method BLASTX
NCBI GI g1353352
BLAST score 494
E value 1.0e-49
Match length 146
% identity 68

NCBI Description (U31975) alanine aminotransferase [Chlamydomonas

reinhardtii]

Seq. No. 2236

Contig ID 1563_8.R1040



```
5'-most EST
                   zzp700834807.h1
Seq. No.
                  2237
Contig ID
                  1563 10.R1040
5'-most EST
                  q4395740
                  2238
Seq. No.
Contig ID
                  1563 11.R1040
5'-most EST
                  asn701138988.h1
Method
                  BLASTN
NCBI GI
                  g16472
BLAST score
                  54
E value
                  2.0e-21
Match length
                  70
% identity
                  94
NCBI Description A.thaliana rRNA repeat unit, most frequent IGR type
Seq. No.
                  2239
Contig ID
                  1565 1.R1040
5'-most EST
                  jC-gmle01810049b11a1
Method
                  BLASTN
NCBI GI
                  g1885325
BLAST score
                  378
E value
                  0.0e+00
Match length
                  786
                  87
% identity
NCBI Description P.sativum mRNA for phosphoribulokinase, partial
Seq. No.
                  2240
                  1565 2.R1040
Contig ID
5'-most EST
                  LIB3055-013-Q1-N1-F6
Method
                  BLASTN
NCBI GI
                  g1885325
BLAST score
                  348
E value
                  0.0e + 00
Match length
                  651
% identity
                  88
NCBI Description P.sativum mRNA for phosphoribulokinase, partial
Seq. No.
                  2241
                  1565 3.R1040
Contig ID
5'-most EST
                  zsg701127883.h1
Method
                  BLASTN
NCBI GI
                  g1885325
BLAST score
                  113
                  2.0e-56
E value
Match length
                  277
                  85
% identity
NCBI Description P.sativum mRNA for phosphoribulokinase, partial
Seq. No.
                  2242
Contig ID
                  1569 1.R1040
5'-most EST
                  LIB3170-055-Q1-K1-C10
```

Method BLASTX NCBI GI g2499497 BLAST score 1898 E value 0.0e + 00



```
Match length
                   406
% identity
                   94
                   PHOSPHOGLYCERATE KINASE, CHLOROPLAST PRECURSOR
NCBI Description
                   >gi 1161600 emb CAA88841 (Z48977) phosphoglycerate kinase
                   [Nicotiana tabacum]
                   2243
Seq. No.
                   1572 1.R1040
Contig ID
5'-most EST
                   LIB3039-006-Q1-E1-H6
Method
                   BLASTX
NCBI GI
                   q2431769
BLAST score
                   199
E value
                   4.0e-15
Match length
                   59
% identity
                   59
NCBI Description
                  (U62752) acidic ribosomal protein Pla [Zea mays]
                   2244
Seq. No.
                   1572 2.R1040
Contig ID
5'-most EST
                   LIB3040-021-Q1-E1-C9
Method
                   BLASTX
NCBI GI
                   g1076627
BLAST score
                   3517
E value
                   0.0e + 00
Match length
                   767
% identity
                   91
                   inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco
NCBI Description
                   >gi_790479_emb_CAA58701_ (X83730) inorganic pyrophosphatase
                   [Nicotiana tabacum]
Seq. No.
                   2245
Contig ID
                   1572 3.R1040
5'-most EST
                   fC-qmse7000754883d1
Method
                  BLASTN
NCBI GI
                   g790478
BLAST score
                   158
E value
                   3.0e-83
Match length
                   434
% identity
                   84
NCBI Description
                  N.tabacum mRNA for inorganic pyrophosphatase (TVP9 clone)
Seq. No.
                   2246
                  1572 4.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy003e03b1
Method
                  BLASTN
```

Method BLASTN
NCBI GI g951322
BLAST score 406
E value 0.0e+00
Match length 689
% identity 90

NCBI Description Vigna radiata pyrophosphatase mRNA, complete cds

Seq. No. 2247

Contig ID 1572_5.R1040

5'-most EST LIB3039-008-Q1-E1-C11

Seq. No. 2248



```
1572 6.R1040
Contig ID
                  LIB3138-015-Q1-N2-D10
5'-most EST
                  BLASTN
Method
                  g951322
NCBI GI
                  90
BLAST score
                  7.0e-43
E value
                  130
Match length
% identity
                  92
NCBI Description Vigna radiata pyrophosphatase mRNA, complete cds
                  2249
Seq. No.
                  1572 7.R1040
Contig ID
                  uaw700664346.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g4249662
BLAST score
                   606
                   4.0e-63
E value
                   134
Match length
                   88
% identity
                  (AF089810) Altered Response to Gravity [Arabidopsis
NCBI Description
                   thaliana]
                   2250
Seq. No.
                   1572 8.R1040
Contig ID
5'-most EST
                  uC-gmronoir002d10b1
                  BLASTN
Method
NCBI GI
                   g2653445
                   182
BLAST score
                   6.0e-98
E value
                   238
Match length
% identity
                   94
NCBI Description Vigna radiata mRNA for proton pyrophosphatase, complete cds
                   2251
Seq. No.
                   1572 10.R1040
Contig ID
5'-most EST
                   q5057881
Method
                   BLASTN
                   g2653445
NCBI GI
                   206
BLAST score
E value
                   1.0e-112
                   510
Match length
                   89
% identity
NCBI Description Vigna radiata mRNA for proton pyrophosphatase, complete cds
Seq. No.
                   2252
                   1572 12.R1040
Contig ID
                   LIB3139-006-P1-N1-G12
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2653445
BLAST score
                   115
E value
                   7.0e-58
                   267
Match length
% identity
NCBI Description Vigna radiata mRNA for proton pyrophosphatase, complete cds
```

2253

1572 14.R1040

Seq. No.

Contig ID



```
epx701105007.h1
5'-most EST
                  BLASTX
Method
                  q485742
NCBI GI
BLAST score
                  674
                  1.0e-145
E value
                  319
Match length
                  85
% identity
                  (L32791) pyrophosphatase [Beta vulgaris]
NCBI Description
Seq. No.
                  2254
                  1574 1.R1040
Contig ID
                  jC-gmst02400073d06a1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q534982
BLAST score
                  595
                  3.0e-61
E value
Match length
                  331
% identity
                  45
                  (X75898) phosphoglucomutase [Spinacia oleracea]
NCBI Description
                  2255
Seq. No.
                  1575 1.R1040
Contig ID
                  LIB3049-043-Q1-E1-A1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g3184053
BLAST score
                  247
                  1.0e-136
E value
Match length
                  423
                  90
% identity
NCBI Description Pisum sativum mRNA for MADS-box transcription factor
                  2256
Seq. No.
                  1575 2.R1040
Contig ID
                  LIB3109-052-Q1-K1-B1
5'-most EST
Method
                  BLASTN
NCBI GI
                   g3184053
BLAST score
                   283
                  1.0e-158
E value
                   559
Match length
% identity
                   88
NCBI Description Pisum sativum mRNA for MADS-box transcription factor
Seq. No.
                   2257
                   1575 3.R1040
Contig ID
5'-most EST
                   asj700967431.hl
                  BLASTN
Method
                   q1239962
NCBI GI
BLAST score
                   116
E value
                   1.0e-58
                   148
Match length
% identity
NCBI Description A.majus mRNA for MADS-box protein (DEFH72)
                   2258
Seq. No.
```

Contig ID 1576_1.R1040

5'-most EST LIB3049-050-Q1-E1-F6

Method BLASTX

NCBI Description



```
q4218169
NCBI GI
                  395
BLAST score
                  4.0e-38
E value
                  120
Match length
% identity
NCBI Description (AJ009724) MADs-box protein, GDEF1 [Gerbera hybrida]
                  2259
Seq. No.
                  1576 2.R1040
Contig ID
                  kl1701207013.h1
5'-most EST
                  BLASTX
Method
                  q4218169
NCBI GI
                   413
BLAST score
                   1.0e-40
E value
                  105
Match length
% identity
NCBI Description (AJ009724) MADs-box protein, GDEF1 [Gerbera hybrida]
                   2260
Seq. No.
                   1577 1.R1040
Contig ID
                   LIB3170-048-Q1-J1-B7
5'-most EST
                   BLASTX
Method
                   q1729860
NCBI GI
BLAST score
                   2053
                   0.0e + 00
E value
                   423
Match length
                   95
% identity
                   26S PROTEASE REGULATORY SUBUNIT 6A HOMOLOG (TAT-BINDING
NCBI Description
                   PROTEIN HOMOLOG 1) (TBP-1) (MG(2+)-DEPENDENT ATPASE 1)
                   (LEMA-1) >gi_1362099_pir__S56672 probable 26S proteinase
                   chain MA-1 - tomato >gi_732815_emb_CAA52445_ (X74426)
                   Mg-dependent ATPase 1 [Lycopersicon esculentum]
                   2261
Seq. No.
                   1577 2.R1040
Contig ID
                   LIB3092-062-Q1-K1-A4
 5'-most EST
                   BLASTX
Method
                   q1777386
NCBI GI
                   853
 BLAST score
                   2.0e-91
 E value
Match length
                   369
 % identity
                   50
                   (U39301) caffeic acid O-methyltransferase [Pinus taeda]
 NCBI Description
                   2262
 Seq. No.
                   1577 3.R1040
 Contig ID
                   cks700764831.h1
 5'-most EST
                   BLASTX
 Method
                   g1729860
 NCBI GI
                   527
 BLAST score
                   9.0e-54
 E value
                   118
 Match length
 % identity
                   87
                   26S PROTEASE REGULATORY SUBUNIT 6A HOMOLOG (TAT-BINDING
```

PROTEIN HOMOLOG 1) (TBP-1) (MG(2+)-DEPENDENT ATPASE 1) (LEMA-1) >gi_1362099_pir__S56672 probable 26S proteinase chain MA-1 - tomato >gi 732815 emb CAA52445 (X74426)

Seq. No.

Contig ID

2268

1580 1.R1040



Mg-dependent ATPase 1 [Lycopersicon esculentum]

2263 Seq. No. Contig ID 1577 4.R1040 5'-most EST jC-qmf102220060e09a1 Method BLASTX q3024434 NCBI GI BLAST score 470 8.0e-70 E value 167 Match length % identity 74 26S PROTEASE REGULATORY SUBUNIT 6A HOMOLOG (TAT-BINDING NCBI Description PROTEIN HOMOLOG 1) (TBP-1) >gi_2564337_dbj_BAA22951_ (D88663) Tat binding protein 1 [Brassica rapa] Seq. No. 1577 5.R1040 Contig ID 5'-most EST zhf700960240.h1 2265 Seq. No. 1577 7.R1040 Contig ID 5'-most EST LIB3028-031-Q1-B1-B9 BLASTX Method g1777386 NCBI GI BLAST score 171 6.0e-12 E value Match length 52 60 % identity (U39301) caffeic acid O-methyltransferase [Pinus taeda] NCBI Description 2266 Seq. No. 1577 8.R1040 Contig ID q4305637 5'-most EST BLASTX Method g3024434 NCBI GI BLAST score 302 2.0e-51 E value 126 Match length 93 % identity 26S PROTEASE REGULATORY SUBUNIT 6A HOMOLOG (TAT-BINDING NCBI Description PROTEIN HOMOLOG 1) (TBP-1) >gi_2564337_dbj_BAA22951_ (D88663) Tat binding protein 1 [Brassica rapa] 2267 Seq. No. 1579 3.R1040 Contig ID jC-gmf102220148g09a15'-most EST BLASTN Method g1870205 NCBI GI BLAST score 265 1.0e-147 E value Match length 576 % identity Medicago sativa (clone nmh7) MADS-box protein mRNA, NCBI Description complete cds



```
5'-most EST
                  LIB3107-052-Q1-K1-B8
Method
                  BLASTX
NCBI GI
                  g2129933
BLAST score
                  691
E value
                  5.0e-73
Match length
                  131
% identity
                  95
NCBI Description
                  myb-related transcription factor TMH27 - tomato
                  >gi 1167484 emb CAA64614 (X95296) transcription factor
                  [Lycopersicon esculentum]
Seq. No.
                  2269
Contig ID
                  1580 2.R1040
5'-most EST
                  hyd700730584.h1
Method
                  BLASTX
NCBI GI
                  q82308
BLAST score
                  412
                  3.0e-40
E value
                  120
Match length
% identity
                  77
NCBI Description myb protein 308 - garden snapdragon
                  2270
Seq. No.
Contig ID
                  1582 1.R1040
5'-most EST
                  leu701152088.h1
Method
                  BLASTX
                  g4105772
NCBI GI
                  301
BLAST score
E value
                  2.0e-27
Match length
                  67
% identity
                  39
NCBI Description
                  (AF049917) PGP9B [Petunia x hybrida]
Seq. No.
                  2271
                  1584_1.R1040
Contig ID
5'-most EST
                  wrg700786265.h2
Method
                  BLASTX
NCBI GI
                  q399024
BLAST score
                  1708
E value
                  0.0e+00
Match length
                  355
% identity
                  93
NCBI Description
                  FRUCTOSE-BISPHOSPHATE ALDOLASE 1, CHLOROPLAST PRECURSOR
                  >gi_322710_pir__S29047 fructose-bisphosphate aldolase (EC
                  4.1.2.13) precursor, chloroplast - garden pea (fragment)
                  >gi 169037 (M97476) aldolase [Pisum sativum]
Seq. No.
                  2272
Contig ID
                  1584 2.R1040
5'-most EST
                  LIB3107-016-Q1-K1-C2
Method
                  BLASTN
                  g169036
NCBI GI
BLAST score
                  189
```

E value 1.0e-102 Match length 369 88 % identity

NCBI Description Pisum sativum L. aldolase gene, 3' end cds

Match length

% identity

193

NCBI Description (AF049917) PGP9B [Petunia x hybrida]

41



```
2273
Seq. No.
                  1584 3.R1040
Contig ID
                  fde700877186.hl
5'-most EST
                  BLASTN
Method
                  q169038
NCBI GI
                  238
BLAST score
                  1.0e-131
E value
                  495
Match length
                  89
% identity
NCBI Description Pisum sativum L. aldolase gene, 3' end cds
                  2274
Seq. No.
                  1584 6.R1040
Contig ID
                  LIB3138-096-Q1-N1-B6
5'-most EST
Method
                  BLASTN
                   q169036
NCBI GI
                   240
BLAST score
                   1.0e-132
E value
                   420
Match length
% identity
                   90
NCBI Description Pisum sativum L. aldolase gene, 3' end cds
                   2275
Seq. No.
                   1586 1.R1040
Contig ID
                   LIB3106-057-Q1-K1-D2
5'-most EST
                   BLASTX
Method
                   g4033424
NCBI GI
                   889
BLAST score
                   8.0e-96
E value
                   208
Match length
                   82
% identity
                   SOLUBLE INORGANIC PYROPHOSPHATASE (PYROPHOSPHATE
NCBI Description
                   PHOSPHO-HYDROLASE) (PPASE) >gi_2668746 (AF034947) inorganic
                   pyrophosphatase [Zea mays]
                   2276
Seq. No.
                   1588 1.R1040
Contig ID
                   pmv7\overline{0}0888137.h1
5'-most EST
                   BLASTX
Method
                   q4105772
NCBI GI
                   818
BLAST score
                   2.0e-87
E value
                   183
Match length
                   54
% identity
                   (AF049917) PGP9B [Petunia x hybrida]
NCBI Description
                   2277
Seq. No.
                   1588 2.R1040
Contig ID
                   LIB3106-005-Q1-K1-H8
5'-most EST
Method
                   BLASTX
                   q4105772
NCBI GI
                   818
BLAST score
                   2.0e-87
E value
```

494



```
2278
Seq. No.
                   1590 1.R1040
Contig ID
                   kmv700737719.h1
5'-most EST
Method
                   BLASTN
                   g2318116
NCBI GI
BLAST score
                   617
                   0.0e + 00
E value
                   1267
Match length
% identity
                   89
NCBI Description Pisum sativum Mg-chelatase subunit D (ChlD) mRNA, complete
                   2279
Seq. No.
                   1591 1.R1040
Contig ID
5'-most EST
                   zhf700960315.h1
Method
                   BLASTN
NCBI GI
                   q1841474
BLAST score
                   171
                   5.0e-91
E value
                   434
Match length
% identity
                   88
NCBI Description P.sativum mRNA for Myb-like protein (Myb26)
                   2280
Seq. No.
                   1591 2.R1040
Contig ID
                   zhf700960664.h1
5'-most EST
                   BLASTN
Method
                   g1841474
NCBI GI
BLAST score
                   227
                   1.0e-124
E value
Match length
                   363
                   91
% identity
NCBI Description P.sativum mRNA for Myb-like protein (Myb26)
                   2281
Seq. No.
                   1592 1.R1040
Contig ID
                   LIB3028-039-Q1-B1-E7
5'-most EST
                   BLASTN
Method
                   g2104680
NCBI GI
                   227
BLAST score
                   1.0e-124
E value
                   894
Match length
                   83
% identity
NCBI Description V.faba mRNA for putative transciption factor (1556bp)
                   2282
 Seq. No.
                   1593 1.R1040
 Contig ID
                   LIB3028-056-Q1-B1-H7
 5'-most EST
                   BLASTN
Method
                   g1223925
NCBI GI
 BLAST score
                   258
                   1.0e-143
E value
                    393
Match length
```

NCBI Description Vigna radiata carboxypeptidase II mRNA, partial cds

92

% identity



Seq. No. 2283

Contig ID 1595_1.R1040

5'-most EST jC-gmf102220055h11a1

Method BLASTX
NCBI GI g2541876
BLAST score 1190
E value 1.0e-131
Match length 465
% identity 53

NCBI Description (D26015) CND41, chloroplast nucleoid DNA binding protein

[Nicotiana tabacum]

Seq. No. 2284

Contig ID 1595 2.R1040

5'-most EST jC-gmle01810023c07d1

Seq. No. 2285

Contig ID 1595_3.R1040 5'-most EST hyd700726241.h1

Method BLASTX
NCBI GI g2541876
BLAST score 306
E value 1.0e-27
Match length 113
% identity 65

NCBI Description (D26015) CND41, chloroplast nucleoid DNA binding protein

[Nicotiana tabacum]

Seq. No. 2286

Contig ID 1595 4.R1040

5'-most EST uC-gmflminsoy115h12b1

Seq. No. 2287

Contig ID 1596_1.R1040 5'-most EST kl1701212428.h1

Method BLASTX
NCBI GI g3979665
BLAST score 269
E value 3.0e-23
Match length 102
% identity 53

NCBI Description (X66482) orf [Arabidopsis thaliana]

Seq. No. 2288

Contig ID 1596_2.R1040 5'-most EST hyd700727018.h1

Seq. No. 2289

Contig ID 1596_3.R1040 5'-most EST gsv701046770.h1

Seq. No. 2290

Contig ID 1596_5.R1040 5'-most EST epx701104051.h1

Seq. No. 2291

Contig ID 1600_1.R1040



```
LIB3028-006-Q1-B1-A7
5'-most EST
Method
                  BLASTX
NCBI GI
                  q4099921
BLAST score
                   451
E value
                  1.0e-44
                  219
Match length
% identity
                   48
NCBI Description
                  (U91982) EREBP-3 homolog [Stylosanthes hamata]
                  2292
Seq. No.
                  1600 2.R1040
Contig ID
5'-most EST
                  LIB3087-005-Q1-K1-D1
Method
                  BLASTN
NCBI GI
                  g4099920
BLAST score
                  95
E value
                  7.0e-46
Match length
                  179
% identity
                   88
NCBI Description Stylosanthes hamata EREBP-3 homolog mRNA, complete cds
                  2293
Seq. No.
Contia ID
                  1601 1.R1040
5'-most EST
                  LIB3028-056-Q1-B1-G1
Method
                  BLASTX
NCBI GI
                   g2370253
BLAST score
                  300
                  3.0e-27
E value
Match length
                  143
% identity
                  51
NCBI Description
                  (Y13273) putative protein kinase [Lycopersicon esculentum]
                   2294
Seq. No.
                  1602 1.R1040
Contig ID
5'-most EST
                   g5057999
Method
                   BLASTX
                   q3860247
NCBI GI
BLAST score
                   1871
                   0.0e+00
E value
Match length
                   385
% identity
                   91
                  (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
                  2295
Seq. No.
                  1602 2.R1040
Contig ID
5'-most EST
                   trc700565474.h1
Method
                  BLASTX
NCBI GI
                   q3860247
BLAST score
                   1863
E value
                   0.0e + 00
Match length
                   383
% identity
                   (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  2296
```

1602 3.R1040

BLASTX

uC-gmronoir011h03b1

Contig ID 5'-most EST

Method



NCBI GI g3860247 BLAST score 539 E value 7.0e-93 Match length 190 % identity 89

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

Seq. No. 2297

Contig ID 1602_4.R1040 5'-most EST rlr700895868.h1

Method BLASTX
NCBI GI g3860247
BLAST score 513
E value 4.0e-52
Match length 114
% identity 84

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

Seq. No. 2298

Contig ID 1602_10.R1040 5'-most EST bth700845653.h1

Method BLASTX
NCBI GI g3860247
BLAST score 329
E value 8.0e-31
Match length 65
% identity 94

NCBI Description (AC005824) unknown protein [Arabidopsis thaliana]

Seq. No. 2299

Contig ID 1604 1.R1040

5'-most EST jC-gmro02910064f06a1

Method BLASTX
NCBI GI g1076389
BLAST score 2747
E value 0.0e+00
Match length 587
% identity 91

NCBI Description protein phosphatase 2A pDF1 - Arabidopsis thaliana

>gi_683502_emb_CAA57528_ (X82002) protein phosphatase 2A 65

kDa regulatory subunit [Arabidopsis thaliana]

Seq. No. 2300

Contig ID 1604_2.R1040

5'-most EST fC-gmse700668369h1

Method BLASTX
NCBI GI g1076389
BLAST score 1095
E value 1.0e-160
Match length 325
% identity 19

NCBI Description protein phosphatase 2A pDF1 - Arabidopsis thaliana

>gi 683502 emb_CAA57528_ (X82002) protein phosphatase 2A 65

kDa regulatory subunit [Arabidopsis thaliana]

Seq. No. 2301

Contig ID 1604 3.R1040



5'-most EST jC-gmle01810047f12d1

Method BLASTX
NCBI GI g1076389
BLAST score 405
E value 2.0e-39
Match length 94
% identity 26

NCBI Description protein phosphatase 2A pDF1 - Arabidopsis thaliana

>gi_683502_emb_CAA57528_ (X82002) protein phosphatase 2A 65

kDa regulatory subunit [Arabidopsis thaliana]

Seq. No. 2302

Contig ID 1604 4.R1040

5'-most EST LIB3170-058-Q1-K1-H1

Method BLASTX
NCBI GI g1568511
BLAST score 299
E value 5.0e-27
Match length 162
% identity 37

NCBI Description (X97913) protein phosphatase 2A [Nicotiana tabacum]

Seq. No. 2303

Contig ID 1604 5.R1040

5'-most EST LIB3109-036-Q1-K1-E3

Method BLASTX
NCBI GI g1076389
BLAST score 326
E value 2.0e-30
Match length 78
% identity 95

NCBI Description protein phosphatase 2A pDF1 - Arabidopsis thaliana

>gi 683502 emb CAA57528 (X82002) protein phosphatase 2A 65

kDa regulatory subunit [Arabidopsis thaliana]

Seq. No. 2304

1606 1.R1040 Contig ID 5'-most EST g5509172 Method BLASTX NCBI GI g1346812 BLAST score 1568 E value 1.0e-175 Match length 418 75 % identity

NCBI Description 26S PROTEASE REGULATORY SUBUNIT 8 (PROTEASOME SUBUNIT P45)

(THYROID HORMONE RECEPTOR INTERACTING PROTEIN 1) (TRIP1) >gi_2136007_pir__I53510 proteasome subunit p45 - human >gi_976227_dbj_BAA07919_ (D44467) 26S proteasome subunit p45 [Homo sapiens] >gi_1096205_prf__2111282A 26S proteasome

[Homo sapiens]

Seq. No. 2305

Contig ID 1606_2.R1040

5'-most EST jC-gmf102220071h07a1

Method BLASTX
NCBI GI g3747111
BLAST score 678



6.0e-71 E value 237 Match length % identity 57

(AF095641) MTN3 homolog [Arabidopsis thaliana] NCBI Description

2306 Seq. No.

Contig ID 1606 4.R1040

5'-most EST LIB3051-112-Q1-K1-B12

BLASTX Method NCBI GI q1171978 BLAST score 1202 E value 1.0e-133 431 Match length 33 % identity

POLYADENYLATE-BINDING PROTEIN 2 (POLY(A) BINDING PROTEIN 2) NCBI Description

(PABP 2) >gi 304109 (L19418) poly(A)-binding protein

[Arabidopsis thaliana] >gi 2911051 emb CAA17561 (AL021961)

poly(A) -binding protein [Arabidopsis thaliana]

2307 Seq. No.

Contig ID 1606 5.R1040

5'-most EST LIB3167-049-P1-K1-G8

BLASTX Method NCBI GI g1171978 BLAST score 1300 E value 1.0e-144 Match length 331 % identity 28

NCBI Description POLYADENYLATE-BINDING PROTEIN 2 (POLY(A) BINDING PROTEIN 2)

(PABP 2) >gi_304109 (L19418) poly(A)-binding protein

[Arabidopsis thaliana] >gi_2911051_emb_CAA17561_ (AL021961)

poly(A)-binding protein [Arabidopsis thaliana]

2308 Seq. No.

1606 6.R1040 Contig ID

5'-most EST LIB3170-049-Q1-J1-B1

Method BLASTX NCBI GI q2213871 BLAST score 295 1.0e-26 E value Match length 73 88 % identity

NCBI Description (AF003126) poly(A)-binding protein [Mesembryanthemum

crystallinum]

2309 Seq. No.

1606 12.R1040 Contig ID 5'-most EST gsv701056823.h1

Method BLASTX g1170767 NCBI GI BLAST score 221 7.0e-18 E value Match length 94 % identity 49

26S PROTEASE REGULATORY SUBUNIT 8 HOMOLOG (LET1 PROTEIN) NCBI Description

>gi_626074_pir__S45176 transcription factor SUG1 homolog - fission yeast (Schizosaccharomyces pombe) >gi_406051



(U02280) Let1 [Schizosaccharomyces pombe] >gi_4106689_emb_CAA22628_ (AL035065) 26s protease regulatory subunit 8 homolog [Schizosaccharomyces pombe]

Seq. No. 2310

Contig ID 1606_13.R1040 5'-most EST ncj700977584.h1

Seq. No. 2311

Contig ID 1606_14.R1040 5'-most EST g4397126

Seq. No. 2312

Contig ID 1606_15.R1040 5'-most EST gsv701044733.h1

Seq. No. 2313

Contig ID 1606_16.R1040

5'-most EST LIB3050-022-Q1-K1-C11

Seq. No. 2314

Contig ID 1606_19.R1040

5'-most EST jC-gmf102220084c12d1

Method BLASTX
NCBI GI g2982331
BLAST score 288
E value 1.0e-25
Match length 58
% identity 98

NCBI Description (AF051251) TAT-binding protein homolog [Picea mariana]

Seq. No. 2315

Contig ID 1607_1.R1040 5'-most EST zpv700757870.h1

Method BLASTX
NCBI GI g2213425
BLAST score 1277
E value 1.0e-141
Match length 279
% identity 84

NCBI Description (Z97064) hypothetical protein [Citrus x paradisi]

Seq. No. 2316

Contig ID 1607_3.R1040

5'-most EST jC-gmro02910073h07a1

Method BLASTN
NCBI GI g2213424
BLAST score 145
E value 1.0e-75
Match length 409
% identity 84

NCBI Description Citrus paradisi mRNA for hypothetical protein

Seq. No. 2317

Contig ID 1607_4.R1040

5'-most EST uC-gmronoir040f06b1



Contig ID 1607_5.R1040

5'-most EST jC-gmro02910005a04d1

Method BLASTX
NCBI GI g2213425
BLAST score 391
E value 7.0e-38
Match length 88
% identity 46

NCBI Description (Z97064) hypothetical protein [Citrus x paradisi]

Seq. No. 2319

Contig ID 1607_6.R1040

5'-most EST LIB3040-059-Q1-E1-B4

Method BLASTX
NCBI GI g2213425
BLAST score 376
E value 3.0e-36
Match length 78
% identity 47

NCBI Description (Z97064) hypothetical protein [Citrus x paradisi]

Seq. No. 2320

Contig ID 1607_9.R1040 5'-most EST zhf700958949.h1

Method BLASTN
NCBI GI g2213424
BLAST score 73
E value 5.0e-33
Match length 177
% identity 86

NCBI Description Citrus paradisi mRNA for hypothetical protein

Seq. No. 2321

Contig ID 1608_1.R1040 5'-most EST uaw700666018.h1

Method BLASTN
NCBI GI g3449326
BLAST score 46
E value 2.0e-16
Match length 151
% identity 86

% identity 86 NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K19M22, complete sequence [Arabidopsis thaliana]

Seq. No. 2322

Contig ID 1608_2.R1040 5'-most EST gsv701049801.h1

Method BLASTN
NCBI GI g3869065
BLAST score 51
E value 2.0e-19
Match length 155
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K24M7, complete sequence [Arabidopsis thaliana]



Contig ID 1611_1.R1040

5'-most EST jC-gmro02910016b07a1

Method BLASTX
NCBI GI g3935150
BLAST score 655
E value 1.0e-68
Match length 194
% identity 45

NCBI Description (AC005106) T25N20.14 [Arabidopsis thaliana]

Seq. No. 2324

Contig ID 1612_1.R1040

5'-most EST LIB3039-046-Q1-E1-H12

Method BLASTN
NCBI GI g563328
BLAST score 116
E value 4.0e-58
Match length 320
% identity 86

NCBI Description A.officinalis mRNA for histone 2B

Seq. No. 2325

Contig ID 1612_2.R1040

5'-most EST jC-gmro02910020f12d1

Method BLASTN
NCBI GI g3510337
BLAST score 34
E value 3.0e-09
Match length 188
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K19E20, complete sequence [Arabidopsis thaliana]

Seq. No. 2326

Contig ID 1613_1.R1040

5'-most EST LIB3051-053-Q1-K2-D4

Seq. No. 2327

Contig ID 1613_3.R1040 5'-most EST zhf700952934.h1

Method BLASTN
NCBI GI g1370171
BLAST score 37
E value 1.0e-11
Match length 70
% identity 89

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB1X

Seq. No. 2328

Contig ID 1614_1.R1040

5'-most EST LIB3106-039-Q1-K1-A10

Method BLASTX
NCBI GI g2253442
BLAST score 268
E value 4.0e-23
Match length 62

503



```
% identity
                   (AF007784) LTCOR11 [Lavatera thuringiaca]
NCBI Description
                   2329
Seq. No.
                   1614 2.R1040
Contig ID
5'-most EST
                   q5126588
                   BLASTX
Method
```

q2253442 NCBI GI 271 BLAST score 9.0e-24 E value 62 Match length 68 % identity

(AF007784) LTCOR11 [Lavatera thuringiaca] NCBI Description

2330 Seq. No.

1614 3.R1040 Contig ID 5'-most EST leu701146240.hl

BLASTX Method q2792297 NCBI GI 175 BLAST score 1.0e-12 E value 48 Match length % identity

(AF039183) GAST-like gene product [Fragaria x ananassa] NCBI Description

2331 Seq. No.

Contig ID 1615 1.R1040

LIB3028-056-Q1-B1-E10 5'-most EST

BLASTX Method NCBI GI q3510538 BLAST score 634 E value 4.0e-66

Match length 124 91 % identity

(U93167) expansin [Prunus armeniaca] NCBI Description

2332 Seq. No.

1615 2.R1040 Contig ID asn701135733.h1 5'-most EST

BLASTX Method g3643603 NCBI GI 586 BLAST score 2.0e-70 E value 192 Match length 68 % identity

(AC005395) unknown protein [Arabidopsis thaliana] NCBI Description

2333 Seq. No.

1617_1.R1040 Contig ID 5'-most EST gsv701055578.hl

BLASTN Method NCBI GI g1262145 BLAST score 201 1.0e-109 E value Match length 533 84 % identity

NCBI Description S.oleracea mRNA for proteasome 37kD subunit



Contig ID 1617_2.R1040 5'-most EST zzp700834205.h1

Method BLASTX
NCBI GI g1709761
BLAST score 433
E value 1.0e-46
Match length 105
% identity 95

NCBI Description PROTEASOME 27 KD SUBUNIT (MULTICATALYTIC ENDOPEPTIDASE COMPLEX 27 KD SUBUNIT) >gi_1262146_emb_CAA65660_ (X96974)

proteasome subunit [Spinacia oleracea]

Seq. No. 2335

Contig ID 1618_1.R1040 5'-most EST pcp700990455.h1

Method BLASTX
NCBI GI g3513727
BLAST score 2263
E value 0.0e+00
Match length 525
% identity 85

NCBI Description (AF080118) contains similarity to TPR domains (Pfam:

TPR.hmm: score: 11.15) and kinesin motor domains (Pfam: kinesin2.hmm, score: 17.49, 20.52 and 10.94) [Arabidopsis thaliana] >gi_4539358_emb_CAB40052.1_ (AL049525) putative

protein [Arabidopsis thaliana]

Seq. No. 2336

Contig ID 1618_2.R1040 5'-most EST zsg701130248.h1

Method BLASTX
NCBI GI g3513727
BLAST score 385
E value 5.0e-37
Match length 148
% identity 42

NCBI Description (AF080118) contains similarity to TPR domains (Pfam:

TPR.hmm: score: 11.15) and kinesin motor domains (Pfam: kinesin2.hmm, score: 17.49, 20.52 and 10.94) [Arabidopsis thaliana] >gi 4539358_emb_CAB40052.1_ (AL049525) putative

protein [Arabidopsis thaliana]

Seq. No. 2337

Contig ID 1620_1.R1040

5'-most EST LIB3094-083-Q1-K1-C2

Seq. No. 2338

Contig ID 1620_2.R1040 5'-most EST leu701152744.h1

Method BLASTX
NCBI GI 94249376
BLAST score 357
E value 3.0e-33
Match length 225
% identity 39

NCBI Description

(AC005966) Strong similarity to gi 2829864 F3I6.6 zinc metalloproteinase homolog from Arabidopsis thaliana BAC gb AC002396. EST gb Z26412 comes from this gene. [Arabidopsis thaliana]

2339 Seq. No.

1620 3.R1040 Contig ID

LIB3056-008-Q1-N1-E3 5'-most EST

BLASTX Method g2827777 NCBI GI 241 BLAST score 5.0e-20 E value 141 Match length % identity 39

METALLOENDOPROTEINASE 1 PRECURSOR (SMEP1) >gi 1679656 NCBI Description

(U63725) metalloproteinase [Glycine max]

2340 Seq. No.

1620 4.R1040 Contig ID

uC-qmflminsoy059d12b1 5'-most EST

2341 Seq. No.

1620 5.R1040 Contig ID

5'-most EST LIB3107-006-Q1-K1-G11

2342 Seq. No.

1620 6.R1040 Contig ID

LIB3051-010-Q1-E1-C4 5'-most EST

BLASTX Method q3776080 NCBI GI 164 BLAST score 4.0e-11 E value 54 Match length 56 % identity

NCBI Description (Y18249) MtN9 [Medicago truncatula]

Seq. No.

2343 1620 7.R1040 Contig ID 5'-most EST leu701148815.hl

Seq. No.

2344 1620 8.R1040 Contig ID uaw700663327.hl 5'-most EST

Seq. No.

2345

1620 10.R1040 Contig ID 5'-most EST uaw700662277.hl

Seq. No.

2346

1621 1.R1040 Contig ID eep700869770.hl 5'-most EST

BLASTX Method g1680686 NCBI GI 497 BLAST score 1.0e-49 E value 189 Match length 53 % identity



```
(U51330) rust resistance kinase Lr10 [Triticum aestivum]
NCBI Description
                  2347
Seq. No.
                  1621 2.R1040
Contig ID
                  LIB3138-008-Q1-N1-A8
5'-most EST
Method
                  BLASTX
                  q3025189
NCBI GI
BLAST score
                  799
                  5.0e-85
E value
                  458
Match length
% identity
                  39
                  HYPOTHETICAL 67.1 KD PROTEIN SLL1770
NCBI Description
                  >gi_1652753_dbj_BAA17672_ (D90908) ABC1-like [Synechocystis
Seq. No.
                   2348
                   1621 3.R1040
Contig ID
                   jC-gmf102220094b06a1
5'-most EST
Seq. No.
                   2349
                   1621 8.R1040
Contig ID
                   jC-gmle01810006e09a1
5'-most EST
                   BLASTX
Method
                   q3025189
NCBI GI
BLAST score
                   192
                   2.0e-14
E value
Match length
                   79
% identity
                   43
                   HYPOTHETICAL 67.1 KD PROTEIN SLL1770
NCBI Description
                   >gi 1652753_dbj_BAA17672_ (D90908) ABC1-like [Synechocystis
                   sp.]
                   2350
Seq. No.
                   1622_1.R1040
Contig ID
                   zhf700961260.hl
5'-most EST
                   BLASTX
Method
                   g2144584
NCBI GI
                   329
BLAST score
                   1.0e-30
E value
                   142
Match length
                   51
% identity
                   trypsin inhibitor A (Kunitz) precursor - soybean
NCBI Description
                   >gi_18770_emb_CAA45777_ (X64447) trypsin inhibitor subtype
                   A [Glycine max]
                   2351
Seq. No.
                   1623 1.R1040
 Contig ID
                   uC-gmrominsoy177g08b1
 5'-most EST
                   BLASTX
Method
                   g1346297
 NCBI GI
 BLAST score
                   1869
                   0.0e + 00
 E value
                   461
 Match length
```

3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE 1 (HMG-COA

hydroxymethylglutaryl-CoA reductase (NADPH) (EC 1.1.1.34) -

REDUCTASE 1) (HMGR1) (HMGR) >gi_2117465_pir__S59944

79

% identity

NCBI Description

potato >gi_169485 (L01400) hydroxymethylglutaryl coenzyme A
reductase [Solanum tuberosum]

2352 Seq. No. 1623 2.R1040 Contig ID zhf700958080.h1 5'-most EST BLASTX Method q1346298 NCBI GI 250 BLAST score 2.0e-21 E value 143 Match length 43 % identity 3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE 2 (HMG-COA NCBI Description REDUCTASE 2) >gi 282943_pir__S25316 hydroxymethylglutaryl-CoA reductase (NADPH) (EC 1.1.1.34) tomato >gi_170452 (M63642) 3-hydroxy-3-methylglutaryl coenzyme A reductase [Lycopersicon esculentum] Seq. No. 2353 1624 1.R1040 Contig ID LIB3028-008-Q1-B1-F1 5'-most EST Method BLASTX q4263048 NCBI GI 800 BLAST score 2.0e-85 E value Match length 271 % identity (AC005142) putative hydrolase [Arabidopsis thaliana] NCBI Description 2354 Seq. No. 1626 1.R1040 Contig ID LIB3028-056-Q1-B1-F10 5'-most EST BLASTX Method q4263791 NCBI GI 289 BLAST score 9.0e-26 E value 200 Match length 35 % identity (AC006068) putative receptor protein kinase [Arabidopsis NCBI Description

Seq. No. 2355

Contig ID 1627_1.R1040 5'-most EST vzy700754339.h1

thaliana]

Seq. No. 2356

Contig ID 1628_1.R1040 5'-most EST jsh701064382.h1

Method BLASTX
NCBI GI g1616612
BLAST score 807
E value 3.0e-86
Match length 193
% identity 79

NCBI Description (Y08423) small GTP-binding protein [Nicotiana

plumbaginifolia]



Contig ID 1628_2.R1040

5'-most EST LIB3040-041-Q1-E1-C9

Method BLASTN
NCBI GI g1184988
BLAST score 108
E value 1.0e-53
Match length 244
% identity 86

NCBI Description Nicotiana tabacum GTP-binding protein NTGB2 mRNA, partial

cds

Seq. No. 2358

Contig ID 1628_3.R1040 5'-most EST asn701139847.h1

Method BLASTX
NCBI GI g266989
BLAST score 533
E value 2.0e-54
Match length 106
% identity 95

NCBI Description GTP-BINDING PROTEIN SAR1B >gi_322517_pir_S28603

GTP-binding protein - Arabidopsis thaliana >gi_166734 (M95795) GTP-binding protein [Arabidopsis thaliana]

Seq. No. 2359

Contig ID 1628_5.R1040 5'-most EST jsh701065569.h1

Method BLASTX
NCBI GI g3334322
BLAST score 308
E value 3.0e-28
Match length 59
% identity 98

NCBI Description GTP-BINDING PROTEIN SAR1B >gi_2108347 (U55036) small

GTP-binding protein Bsarlb [Brassica rapa]

Seq. No. 2360

Contig ID 1629_1.R1040

5'-most EST LIB3049-022-Q1-E1-G4

Method BLASTN
NCBI GI g1370286
BLAST score 173
E value 4.0e-92
Match length 385
% identity 86

NCBI Description P.sativum mRNA for pore protein

Seq. No. 2361

Contig ID 1629_2.R1040 5'-most EST hrw701060880.h1

Seq. No. 2362

Contig ID 1629_3.R1040

5'-most EST LIB3109-033-Q1-K1-A4

Method BLASTX NCBI GI g1370287

BLAST score 211 E value 2.0e-16 Match length 52 % identity 79

NCBI Description (Z73553) core protein [Pisum sativum]

Seq. No.

Contig ID 1630_1.R1040

5'-most EST LIB3028-056-Q1-B1-C1

2363

2364

Seq. No.

Contig ID 1631_1.R1040

5'-most EST jC-gmf102220050d12a1

Seq. No. 2365

Contig ID 1633 1.R1040

5'-most EST LIB3087-011-Q1-K1-F10

Method BLASTN
NCBI GI g510189
BLAST score 445
E value 0.0e+00
Match length 815
% identity 89

NCBI Description P.sativum (miranda) mRNA for chloroplast outer envelope

protein

Seq. No. 2366

Contig ID 1633 3.R1040 5'-most EST leu701144743.h1

Method BLASTN
NCBI GI g3293200
BLAST score 201
E value 1.0e-109
Match length 305
% identity 91

NCBI Description Pisum sativum GTP-binding protein (IAP34) mRNA, complete

cds

Seq. No. 2367

Contig ID 1634_1.R1040

5'-most EST uC-gmropic011a08b1

Method BLASTX
NCBI GI g2191171
BLAST score 587
E value 2.0e-60
Match length 191
% identity 60

NCBI Description (AF007270) similar to A. thaliana DI19 mRNA (NID:g469110)

[Arabidopsis thaliana]

Seq. No. 2368

Contig ID 1636_1.R1040

5'-most EST jC-gmf102220053f02a1

Method BLASTX
NCBI GI g2239262
BLAST score 576
E value 3.0e-59



Match length 212 % identity 51

% identity 51
NCBI Description (Y1)

(Y13285) pectin methylesterase-like protein [Zea mays]

Seq. No. 2369

Contig ID 1637_1.R1040

5'-most EST fC-qmse700754179d2

Method BLASTX
NCBI GI g2961107
BLAST score 245
E value 2.0e-20
Match length 91
% identity 49

NCBI Description (AF042383) TLS-associated protein with SR repeats [Mus

musculus] >gi_2961149 (AF047448) TLS-associated protein

TASR [Homo sapiens]

Seq. No. 2370

Contig ID 1637 2.R1040

5'-most EST LIB3028-056-Q1-B1-C7

Seq. No. 2371

Contig ID 1638_1.R1040

5'-most EST uC-gmflminsoy069e04b1

Method BLASTX
NCBI GI g2959767
BLAST score 2370
E value 0.0e+00
Match length 662
% identity 47

NCBI Description (AJ002584) AtMRP4 [Arabidopsis thaliana] >gi_3738292

(AC005309) glutathione-conjugate transporter AtMRP4

[Arabidopsis thaliana]

Seq. No. 2372

Contig ID 1639 1.R1040

5'-most EST uC-qmflminsoy011c12b1

Method BLASTX
NCBI GI g886132
BLAST score 299
E value 4.0e-27
Match length 135
% identity 47

NCBI Description (U28149) putative surface protein [Medicago sativa]

Seq. No. 2373

Contig ID 1640_1.R1040 5'-most EST epx701109865.h1

Method BLASTN
NCBI GI g391604
BLAST score 341
E value 0.0e+00
Match length 953
% identity 84

NCBI Description Arabidopsis thaliana ATPK15D mRNA for casein kinase II

catalytic subunit, complete cds



 Seq. No.
 2374

 Contig ID
 1640_2.R1040

 5'-most EST
 wrg700790238.h2

Method BLASTN
NCBI GI g391604
BLAST score 134
E value 6.0e-69
Match length 350
% identity 85

% identity 85
NCBI Description Arabidopsis thaliana ATPK15D mRNA for casein kinase II

catalytic subunit, complete cds

Seq. No. 2375

Contig ID 1640_4.R1040

5'-most EST jC-gmle01810021e08a1

Method BLASTN
NCBI GI g3169169
BLAST score 47
E value 5.0e-17
Match length 95
% identity 53

NCBI Description Arabidopsis thaliana chromosome II BAC F21P24 genomic

sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 2376

1640 5.R1040 Contig ID g5677543 5'-most EST BLASTX Method q585350 NCBI GI BLAST score 730 2.0e-77 E value 144 Match length 94 % identity

NCBI Description CASEIN KINASE II, ALPHA CHAIN 2 (CK II)

>gi_419753_pir__S31099 casein kinase II (EC 2.7.1.-)
alpha-type chain (clone ATCKA2) - Arabidopsis thaliana
>gi 391605_dbj_BAA01091_ (D10247) casein kinase II

catalytic subunit [Arabidopsis thaliana]

Seq. No. 2377

Contig ID 1642_1.R1040 5'-most EST jsh701067642.h1

Method BLASTX
NCBI GI g4056468
BLAST score 359
E value 8.0e-34
Match length 154
% identity 46

NCBI Description (AC005990) Contains similarity to gb_X66426

polygalacturonase from Persea americana and is a member of the signal peptidase family PF_00461 and polygalacturonase

family PF_00295. [Arabidopsis thaliana]

Seq. No. 2378

Contig ID 1643_1.R1040 5'-most EST g5509675 Method BLASTX



g2497492 NCBI GI BLAST score 364 2.0e-34 E value Match length 101 % identity 68

URIDYLATE KINASE (UK) (URIDINE MONOPHOSPHATE KINASE) (UMP NCBI Description

KINASE) >gi 1653646 dbj_BAA18558_ (D90915) uridine

monophosphate kinase [Synechocystis sp.]

2379 Seq. No.

1644 1.R1040 Contig ID g5677499 5'-most EST BLASTX Method q4539400 NCBI GI 824 BLAST score 2.0e-88 E value 208 Match length

% identity (AL035526) putative protein [Arabidopsis thaliana] NCBI Description

2380 Seq. No.

1645 1.R1040 Contig ID pmv700893145.h15'-most EST

75

BLASTX Method g1723242 NCBI GI 152 BLAST score 2.0e-09 E value 74 Match length 50 % identity

HYPOTHETICAL 36.8 KD PROTEIN C26A3.16 IN CHROMOSOME I NCBI Description >gi 1177363 emb CAA93239 (Z69240) yeast dsk2 homolog,

ubiquitin-like protein [Schizosaccharomyces pombe]

2381 Seq. No.

1646_1.R1040 Contig ID $qsv7\overline{0}1051466.h1$ 5'-most EST

BLASTX Method NCBI GI q4510345 355 BLAST score E value 4.0e-33 Match length 164 45 % identity

(AC006921) unknown protein [Arabidopsis thaliana] NCBI Description

2382 Seq. No.

1646 2.R1040 Contig ID 5'-most EST sat701008658.hl

Seq. No.

2383

1646 3.R1040 Contig ID

LIB3170-007-Q1-K1-G8 5'-most EST

Seq. No.

2384

1646 4.R1040 Contig ID

5'-most EST LIB3051-032-Q1-K1-F1

2385

Seq. No.



1646 5.R1040 Contig ID pxt700945341.h1 5'-most EST

Method BLASTX NCBI GI g4510345 BLAST score 145 5.0e-09 E value Match length 36 % identity 67

(AC006921) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

1646 6.R1040 Contig ID

LIB3138-021-Q1-N1-H12 5'-most EST

2386

Seq. No. 2387

Contig ID 1646 8.R1040

LIB3139-090-P1-N1-H4 5'-most EST

Method BLASTX NCBI GI g1172650 BLAST score 164 E value 6.0e-11 Match length 74 % identity 47

EXTRACELLULAR METALLOPROTEASE PRECURSOR NCBI Description

>gi_95576_pir_ A41048 extracellular metalloproteinase (EC 3.4.24.-) precursor - Erwinia carotovora subsp. carotovora >gi_148480 (M36651) extracellular protease (prt) [Erwinia

carotovora]

2388 Seq. No.

1646 11.R1040 Contig ID

5'-most EST jC-qmf102220148b01d1

2389 Seq. No.

Contig ID 1646 12.R1040 bth700844489.h1 5'-most EST

BLASTX Method NCBI GI g3123056 BLAST score 152 1.0e-09 E value Match length 174 33 % identity

HYPOTHETICAL PROTEIN KIAA0039 >gi 436222_dbj_BAA05039_ NCBI Description

(D26018) KIAA0039 [Homo sapiens]

Seq. No. 2390

1646 17.R1040 Contig ID

5'-most EST LIB3\overline{1}38-001-Q1-N1-A12

Seq. No.

Contig ID 1646 18.R1040 5'-most EST fua701037588.hl

2392 Seq. No.

Contig ID 1647 1.R1040

5'-most EST LIB3093-038-Q1-K1-A3

Method BLASTN



NCBI GI g488738 BLAST score 233 E value 1.0e-128 Match length 777 % identity 82

NCBI Description G.hirsutum (DPL 62) mRNA for ribosomal protein small

subunit 4e

Seq. No. 2393

Contig ID 1647_2.R1040 5'-most EST pmv700890611.h1

Method BLASTX
NCBI GI g2262167
BLAST score 629
E value 1.0e-65
Match length 170
% identity 72

NCBI Description (AC002329) cytosolic ribosomal protein S4 [Arabidopsis

thaliana]

Seq. No. 2394

Contig ID 1648_1.R1040 5'-most EST zhf700963763.h1

Method BLASTX
NCBI GI g2827628
BLAST score 268
E value 2.0e-23
Match length 122
% identity 45

NCBI Description (AL021636) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2395

Contig ID 1650 1.R1040

5'-most EST LIB3138-126-Q1-N1-B10

Method BLASTN
NCBI GI g1853967
BLAST score 181
E value 8.0e-97
Match length 760
% identity 81

NCBI Description Vigna unguiculata mRNA for CPRD12 protein, complete cds

Seq. No. 2396

Contig ID 1650_2.R1040

5'-most EST LIB3170-041-Q1-J1-A11

Method BLASTN
NCBI GI g1853967
BLAST score 66
E value 1.0e-28
Match length 138
% identity 87

NCBI Description Vigna unguiculata mRNA for CPRD12 protein, complete cds

Seq. No. 2397

Contig ID 1651_1.R1040 5'-most EST gsv701054270.h1

Method BLASTX



NCBI GI g114181 BLAST score 915 E value 1.0e-98 Match length 324 % identity 57

NCBI Description

3-DEHYDROQUINATE SYNTHASE >gi_68385_pir__SYECQ
3-dehydroquinate synthase (EC 4.6.1.3) - Escherichia coli
>gi_40968_emb_CAA27495_ (X03867) 3-dehydroquinate synthase
(aa 1-362) [Escherichia coli] >gi_41225_emb_CAA79666_
(Z19601) ORF, aroB. Millar G., Coggins J.R.; FEBS Lett.
200:11-17(1986) [Escherichia coli] >gi_606323 (U18997)
3-dehydroquinate synthase [Escherichia coli] >gi_1789791
(AE000414) 3-dehydroquinate synthase [Escherichia coli]

Seq. No. 2398 Contig ID 1651_2.R1040

5'-most EST LIB3138-082-P1-N1-E1

Seq. No. 2399

Contig ID 1651_3.R1040

5'-most EST fC-gmle700875818f1

Seq. No. 2400

Contig ID 1652_1.R1040

5'-most EST LIB3139-081-P1-N1-F5

Seq. No. 2401

Contig ID 1652 2.R1040

5'-most EST LIB3092-015-Q1-K1-E7

Seq. No. 2402

Contig ID 1652_3.R1040

5'-most EST LIB3092-033-Q1-K1-H10

Seq. No. 2403

Contig ID 1653_1.R1040

5'-most EST LIB3138-129-Q1-N1-E3

Method BLASTX
NCBI GI g2465008
BLAST score 1115
E value 1.0e-122
Match length 339
% identity 65

NCBI Description (AJ001445) ripening-induced protein [Fragaria vesca]

Seq. No. 2404

Contig ID 1654 1.R1040 5'-most EST zhf700962965.h1

Method BLASTX
NCBI GI g1076287
BLAST score 1503
E value 1.0e-167
Match length 419
% identity 67

NCBI Description amine acid permease - Arabidopsis thaliana

>qi 510236 emb CAA50672 (X71787) amine acid permease

[Arabidopsis thaliana]



```
2405
Seq. No.
                   1654 2.R1040
Contig ID
                   V4R-\overline{0}1-Q1-B1-B10
5'-most EST
Method
                   BLASTX
                   g1076287
NCBI GI
                   155
BLAST score
                   2.0e-10
E value
                   56
Match length
                   57
% identity
                   amine acid permease - Arabidopsis thaliana
NCBI Description
                   >gi 510236_emb_CAA50672_ (X71787) amine acid permease
                    [Arabidopsis thaliana]
Seq. No.
                   2406
                   1656 1.R1040
Contig ID
                   ncj7\overline{0}0980372.h1
5'-most EST
                   BLASTX
Method
                    q4056480
NCBI GI
BLAST score
                    724
                    3.0e-91
E value
                    224
Match length
                    76
% identity
                    (AC005896) putative adenylate kinase [Arabidopsis thaliana]
NCBI Description
                    2407
Seq. No.
                    1656 2.R1040
Contig ID
                   LIB3170-024-Q1-K1-C5
5'-most EST
                    BLASTX
Method
                    g4056480
NCBI GI
                    453
BLAST score
                    4.0e-45
E value
                    113
Match length
                    81
 % identity
                    (AC005896) putative adenylate kinase [Arabidopsis thaliana]
NCBI Description
                    2408
 Seq. No.
                    1656 3.R1040
 Contig ID
                    uC-gmropic101b11b1
 5'-most EST
 Method
                    BLASTX
                    q4056480
 NCBI GI
                    417
 BLAST score
                    7.0e-41
 E value
                    146
 Match length>
 % identity
                    62
                    (AC005896) putative adenylate kinase [Arabidopsis thaliana]
 NCBI Description
                    2409
 Seq. No.
 Contig ID
                    1656 4.R1040
 5'-most EST
                    sat701006284.h2
 Method
                    BLASTX
 NCBI GI
                    q4056480
                    227
 BLAST score
                    1.0e-18
 E value
 Match length
                    82
```

(AC005896) putative adenylate kinase [Arabidopsis thaliana]

63

% identity

NCBI Description



Contig ID 1657_1.R1040

5'-most EST LIB3028-004-Q1-B1-G5

Seq. No. 2411

Contig ID 1658 1.R1040

5'-most EST LIB3028-056-Q1-B1-B7

Method BLASTX
NCBI GI g4049341
BLAST score 380
E value 2.0e-36
Match length 175
% identity 41

NCBI Description (AL034567) putative protein [Arabidopsis thaliana]

Seq. No. 2412

Contig ID 1659_1.R1040

5'-most EST LIB3\(\bar{1}\)06-041-Q1-K1-E9

Method BLASTX
NCBI GI g4335745
BLAST score 618
E value 5.0e-64
Match length 270
% identity 47

NCBI Description (AC006284) putative hydrolase (contains an

esterase/lipase/thioesterase active site serine domain

(prosite: PS50187) [Arabidopsis thaliana]

Seq. No. 2413

Contig ID 1659_2.R1040 5'-most EST ncj700981794.h1

Method BLASTX
NCBI GI g4335745
BLAST score 362
E value 3.0e-34
Match length 141
% identity 49

NCBI Description (AC006284) putative hydrolase (contains an

esterase/lipase/thioesterase active site serine domain

(prosite: PS50187) [Arabidopsis thaliana]

Seq. No. 2414

Contig ID 1659_3.R1040 5'-most EST ssr700560096.h1

Seq. No. 2415

11.23 5 1.

Contig ID 1660 1.R1040

5'-most EST jC-gmro02910040d04a1

Method BLASTX
NCBI GI g4309757
BLAST score 229
E value 1.0e-18
Match length 63
% identity 67

NCBI Description (AC006217) hypothetical protein [Arabidopsis thaliana]



```
Seq. No.
                  1665 1.R1040
Contig ID
5'-most EST
                  LIB3170-046-Q1-J1-G5
Method
                  BLASTX
                  g2239262
NCBI GI
BLAST score
                  1126
                  1.0e-123
E value
                   465
Match length
% identity
                   47
                  (Y13285) pectin methylesterase-like protein [Zea mays]
NCBI Description
                  2417
Seq. No.
                  1665 2.R1040
Contig ID
5'-most EST
                   jC-gmf102220086f12d1
                   2418
Seq. No.
                   1667 1.R1040
Contig ID
                   jC-gmro02910052a03a1
5'-most EST
                   BLASTX
Method
                   q4105798
NCBI GI
BLAST score
                   1120
                   1.0e-122
E value
                   270
Match length
                   65
% identity
                  (AF049930) PGP237-11 [Petunia x hybrida]
NCBI Description
                   2419
Seq. No.
                   1667 2.R1040
Contig ID
5'-most EST
                  LIB3051-067-Q1-K1-D8
Seq. No.
                   2420
                   1667 3.R1040
Contig ID
5'-most EST
                   leu7\overline{0}1156628.h1
Seq. No.
                   2421
Contig ID
                   1667 4.R1040
5'-most EST
                   LIB3107-050-Q1-K1-F1
                   2422
Seq. No.
Contig ID
                   1667 5.R1040
                   g4300883
5'-most EST
Method
                   BLASTX
                   g4105798
NCBI GI
BLAST score
                   416
E value
                   8.0e-57
                   171
Match length
% identity
                   56
NCBI Description (AF049930) PGP237-11 [Petunia x hybrida]
                   2423
Seq. No.
```

Contig ID 1667 6.R1040

5'-most EST jC-qmf102220063d05d1

Method BLASTX NCBI GI g4105798 BLAST score 201 E value 1.0e-15 Match length 41



% identity 83

NCBI Description (AF049930) PGP237-11 [Petunia x hybrida]

Seq. No. 2424

Contig ID 1669_1.R1040 5'-most EST eep700865012.h1

Method BLASTX
NCBI GI g2194132
BLAST score 1103
E value 1.0e-120
Match length 355
% identity 66

NCBI Description (AC002062) No definition line found [Arabidopsis thaliana]

Seq. No. 2425

Contig ID 1670 1.R1040

5'-most EST LIB3139-116-P1-N1-E11

Method BLASTX
NCBI GI g3763918
BLAST score 701
E value 6.0e-74
Match length 194
% identity 69

NCBI Description (AC004450) putative isopropylmalate dehydratase

[Arabidopsis thaliana]

Seq. No. 2426

Contig ID 1670_2.R1040

5'-most EST LIB3049-027-Q1-E1-D3

Method BLASTX
NCBI GI g3763918
BLAST score 214
E value 3.0e-17
Match length 48
% identity 81

NCBI Description (AC004450) putative isopropylmalate dehydratase

[Arabidopsis thaliana]

Seq. No. 2427

Contig ID 1672_1.R1040 5'-most EST wrg700788550.h1

Method BLASTX
NCBI GI g2262116
BLAST score 218
E value 1.0e-17
Match length 120
% identity 39

NCBI Description (AC002343) cellulose synthase isolog [Arabidopsis thaliana]

Seq. No. 2428

Contig ID 1673_1.R1040 5'-most EST leu701150051.h1

Seq. No. 2429

Contig ID 1673_2.R1040

5'-most EST jC-gmfl02220127d06d1



```
Seq. No.
                   1677 1.R1040
Contig ID
5'-most EST
                   LIB3170-050-Q1-K1-H7
Seq. No.
                   2431
                   1677 2.R1040
Contig ID
5'-most EST
                   leu701153574.h1
Method
                   BLASTX
NCBI GI
                   g3201613
BLAST score
                   658
E value
                   1.0e-68
                   214
Match length
% identity
                   57
                  (AC004669) glutathione S-transferase [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                   1677_3.R1040
5'-most EST
                   uC-gmropic116g02b1
Seq. No.
                   2433
Contig ID
                   1678 1.R1040
5'-most EST
                   kmv700743274.h1
Method
                   BLASTX
                   g2582822
NCBI GI
BLAST score
                   956
E value
                   1.0e-103
                   290
Match length
% identity
                   67
                  (Y09987) CDSP32 protein (Chloroplast Drought-induced Stress
NCBI Description
                   Protein of 32kDa) [Solanum tuberosum]
                   2434
Seq. No.
                   1679 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220051g10a1
Method
                   BLASTX
NCBI GI
                   g2454182
BLAST score
                   1691
                   0.0e+00
E value
Match length
                   395
                   83
% identity
                   (U80185) pyruvate dehydrogenase E1 alpha subunit
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   2435
Contig ID
                   1685 1.R1040
5'-most EST
                   jsh701065040.h1
Method
                   BLASTN
NCBI GI
                   q3241927
```

BLAST score 46 2.0e-16 E value Match length 114 % identity 85

Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone: NCBI Description

MTE17, complete sequence [Arabidopsis thaliana]

Seq. No. 2436

Contig ID 1686 1.R1040



165

5'-most EST epx701105991.h1
Method BLASTX
NCBI GI g542100
BLAST score 718
E value 6.0e-76

% identity 82 NCBI Description H-

Match length

H-protein - Flaveria pringlei >gi_2147453_pir__S60195 H-protein precursor (clone HFC1) - Flaveria cronquistii >gi_437993_emb_CAA81073_(Z25854) H-protein [Flaveria cronquistii] >gi_437999_emb_CAA81074_(Z25855) H-protein [Flaveria pringlei] >gi_3688299_emb_CAB16912_(Z99763)

H-protein [Flaveria pringlei]

Seq. No. 2437

Contig ID 1688 1.R1040

5'-most EST jC-gmst02400047c03a1

Method BLASTX
NCBI GI g113116
BLAST score 1475
E value 1.0e-164
Match length 503
% identity 55

NCBI Description ATP-CITRATE (PRO-S-)-LYASE (CITRATE CLEAVAGE ENZYME)

>gi_111396_pir__A35007 ATP citrate (pro-S)-lyase (EC
4.1.3.8) - rat >gi_203490 (J05210) ATP citrate-lyase

[Rattus norvegicus]

Seq. No. 2438

Contig ID 1688_2.R1040

5'-most EST LIB3170-040-Q1-K1-F8

Method BLASTX
NCBI GI g4406759
BLAST score 430
E value 5.0e-42
Match length 218
% identity 43

NCBI Description (AC006836) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2439

Contig ID 1688 3.R1040

5'-most EST uC-gmrominsoy221a03b1

Method BLASTX
NCBI GI g4107343
BLAST score 347
E value 1.0e-32
Match length 160
% identity 43

NCBI Description (AJ224922) ATP citrate lyase [Sordaria macrospora]

Seq. No.

Contig ID 1688 5.R1040

5'-most EST LIB3051-001-Q1-E1-D1

2440

Seq. No. 2441

Contig ID 1688 9.R1040

5'-most EST uC-gmrominsoy177g10b1

```
BLASTX
Method
                  g4107343
NCBI GI
                  140
BLAST score
                  8.0e-09
E value
                  61
Match length
% identity
                   48
                  (AJ224922) ATP citrate lyase [Sordaria macrospora]
NCBI Description
                   2442
Seq. No.
                  1688 12.R1040
Contig ID
5'-most EST
                  uxk700667645.h1
                  BLASTX
Method
                   q1703083
NCBI GI
                   327
BLAST score
                   3.0e-30
E value
Match length
                   103
% identity
                  ATP-CITRATE (PRO-S-)-LYASE (CITRATE CLEAVAGE ENZYME)
NCBI Description
                   >gi 105392 pir __S21173 ATP citrate (pro-S)-lyase - human
                   >gi_28935_emb_CAA45614_ (X64330) ATP-citrate (pro-S-)-lyase
                   [Homo sapiens] >gi_4501865_ref_NP_001087.1_pACLY_ ATP
                   citrate lyase
                   2443
Seq. No.
                   1689 1.R1040
Contig ID
                   LIB3028-054-Q1-B1-F8
5'-most EST
Method
                   BLASTX
                   q1346701
NCBI GI
BLAST score
                   418
E value
                   6.0e-41
                   164
Match length
                   51
% identity
                  EXOPOLYGALACTURONASE CLONE GBGE184 PRECURSOR (EXOPG)
NCBI Description
                   (PECTINASE) (GALACTURAN 1,4-ALPHA-GALACTURONIDASE)
                   >gi_421831_pir__S34199 exopolygalacturonase (clone GBGe184)
                   - Arabidopsis thaliana >gi_313682_emb_CAA51032_ (X72291)
                   exopolygalacturonase [Arabidopsis thaliana]
                   >gi_3004440_emb_CAA76127_ (Y16230) polygalacturonase
                   [Arabidopsis thaliana]
```

2444 Seq. No. 1690 1.R1040 Contig ID zhf700964274.h1 5'-most EST BLASTX Method

g2827554 NCBI GI 199 BLAST score 4.0e-15 E value 149 Match length % identity

(AL021635) putative DNA binding protein [Arabidopsis NCBI Description

thaliana]

2445 Seq. No.

1690 2.R1040 Contig ID

uC-qmropic029d08b1 5'-most EST

2446 Seq. No.



```
1690 3.R1040
Contig ID
                  uC-gmflminsoy030a12b1
5'-most EST
Seq. No.
                  2447
                  1690 4.R1040
Contig ID
                  gsv701050564.h1
5'-most EST
                  2448
Seq. No.
                  1694 1.R1040
Contig ID
                   awf7\overline{0}0840660.h1
5'-most EST
                  BLASTX
Method
                   g3738333
NCBI GI
                   601
BLAST score
                   6.0e-62
E value
Match length
                   186
% identity
                   65
                  (AC005170) unknown protein [Arabidopsis thaliana]
NCBI Description
                   2449
Seq. No.
                   1695 1.R1040
Contig ID
                   jC-gmf102220142a10a1
5'-most EST
                   BLASTX
Method
                   g4314378
NCBI GI
                   277
BLAST score
                   2.0e-24
E value
                   101
Match length
                   59
% identity
                   (AC006232) putative lipase [Arabidopsis thaliana]
NCBI Description
                   2450
Seq. No.
                   1696 1.R1040
Contig ID
                   LIB3028-054-Q1-B1-E6
5'-most EST
                   2451
Seq. No.
                   1699 1.R1040
Contig ID
                   pcp700993505.h1
5'-most EST
                   BLASTN
Method
                   g3449322
NCBI GI
                   42
BLAST score
                   3.0e-14
E value
Match length
                   389
                   82
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:
NCBI Description
                   MXC17, complete sequence [Arabidopsis thaliana]
                   2452
Seq. No.
                   1700 1.R1040
Contig ID
                   leu701154008.h1
5'-most EST
Method
                   BLASTN
                   q1420884
NCBI GI
```

Method BLASTN
NCBI GI g1420884
BLAST score 186
E value 1.0e-100
Match length 481
% identity 87

NCBI Description Phaseolus vulgaris proline-rich 14 kDa protein mRNA,

complete cds



```
2453
Seq. No.
                   1700 2.R1040
Contig ID
                   q4298459
5'-most EST
                   BLASTX
Method
NCBI GI
                   q462195
                    476
BLAST score
                    1.0e-47
E value
                    112
Match length
% identity
                    PROTEIN TRANSLATION FACTOR SUI1 HOMOLOG (GOS2 PROTEIN)
NCBI Description
                    >gi_100682_pir__S21636 GOS2 protein - rice
                    >gi_20238 emb_CAA36190_ (X51910) GOS2 [Oryza sativa] >gi_3789950 (AF094774) translation initiation factor [Oryza
                    sativa]
Seq. No.
                    2454
                    1700 3.R1040
Contig ID
                    fde700877279.h1
5'-most EST
                    BLASTX
Method
                    g3659907
NCBI GI
BLAST score
                    429
                    4.0e-42
E value
                    107
Match length
                    79
% identity
                    (AF091857) protein translation factor SUI1 homolog
NCBI Description
                    [Pimpinella brachycarpa]
                    2455
Seq. No.
                    1700 4.R1040
Contig ID
                    vwf7\overline{0}0674561.h1
5'-most EST
                    BLASTN
Method
                    g1420884
NCBI GI
                    107
BLAST score
                    5.0e-53
E value
                    347
Match length
                    88
 % identity
                    Phaseolus vulgaris proline-rich 14 kDa protein mRNA,
NCBI Description
                    complete cds
                    2456
 Seq. No.
                    1700 5.R1040
 Contig ID
                    vzy700753816.h1
 5'-most EST
                    BLASTX
Method
                    g4140371
 NCBI GI
                     402
 BLAST score
                    5.0e-39
 E value
                     180
 Match length
                     53
 % identity
                    (U43711) 3-hydroxy-3-methylglutaryl-coenzyme A reductase
 NCBI Description
                     [Morus alba]
```

1700 6.R1040 Contig ID jsh701063980.h15'-most EST

BLASTN Method g1420884 NCBI GI BLAST score 152

1.0e-79 E value 379 Match length 87 % identity

Phaseolus vulgaris proline-rich 14 kDa protein mRNA, NCBI Description

complete cds

2458 Seq. No.

1700 8.R1040 Contig ID 5'-most EST g4313847 Method BLASTX g462195 NCBI GI 368 BLAST score 6.0e-35 E value 88 Match length

% identity PROTEIN TRANSLATION FACTOR SUI1 HOMOLOG (GOS2 PROTEIN) NCBI Description

>gi 100682 pir S21636 GOS2 protein - rice

>gi_20238_emb_CAA36190_ (X51910) GOS2 [Oryza sativa]

>gi 3789950 (AF094774) translation initiation factor [Oryza

sativa]

81

2459 Seq. No.

1700 12.R1040 Contig ID hrw701059318.h1 5'-most EST

BLASTX Method g3659907 NCBI GI 167 BLAST score 6.0e-12 E value Match length 38 79 % identity

(AF091857) protein translation factor SUI1 homolog NCBI Description

[Pimpinella brachycarpa]

2460 Seq. No.

1701 1.R1040 Contig ID

5'-most EST LIB3109-037-Q1-K1-E3

BLASTX Method q3183088 NCBI GI 316 BLAST score 1.0e-28 E value 96 Match length 56 % identity

PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR NCBI Description

(LTP) >gi_629658_pir__S47084 lipid transfer like protein cowpea >gi 499034 emb CAA56113 (X79604) lipid transfer

like protein [Vigna unguiculata]

Seq. No. 2461

1701 2.R1040 Contig ID

LIB3107-040-Q1-K1-B8 5'-most EST

Method BLASTX q3183088 NCBI GI 309 BLAST score 6.0e-28 E value Match length 96 59 % identity

NCBI Description PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR



(LTP) >gi_629658_pir__S47084 lipid transfer like protein -cowpea >gi_499034_emb_CAA56113_ (X79604) lipid transfer like protein [Vigna unguiculata]

Seq. No. 2462

Contig ID 1701 3.R1040

5'-most EST LIB3109-044-Q1-K1-G3

Method BLASTX
NCBI GI g3183088
BLAST score 208
E value 4.0e-24
Match length 96
% identity 56

NCBI Description PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR

(LTP) >gi_629658_pir__S47084 lipid transfer like protein - cowpea >gi_499034 emb_CAA56113_ (X79604) lipid transfer

like protein [Vigna unguiculata]

Seq. No. 2463

Contig ID 1701 4.R1040

5'-most EST LIB3106-063-Q1-K1-G7

Method BLASTX
NCBI GI g3183088
BLAST score 316
E value 9.0e-29
Match length 96
% identity 56

NCBI Description PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR

(LTP) >gi_629658_pir__S47084 lipid transfer like protein -cowpea >gi_499034 emb_CAA56113 (X79604) lipid transfer

like protein [Vigna unguiculata]

Seq. No. 2464

Contig ID 1702_1.R1040

5'-most EST LIB3072-057-Q1-K1-C5

Method BLASTX
NCBI GI g3168840
BLAST score 292
E value 4.0e-26
Match length 67
% identity 79

NCBI Description (U88711) copper homeostasis factor [Arabidopsis thaliana]

Seq. No. 2465

Contig ID 1702_3.R1040

5'-most EST LIB3040-036-Q1-E1-C2

Method BLASTX
NCBI GI g3168840
BLAST score 292
E value 3.0e-26
Match length 67
% identity 79

NCBI Description (U88711) copper homeostasis factor [Arabidopsis thaliana]

Seq. No.

2466

Contig ID 1703_1.R1040 5'-most EST ncj700982230.h1



BLASTX Method g2244792 NCBI GI 973 BLAST score 1.0e-106 E value 225 Match length 77 % identity (Z97336) ankyrin homolog [Arabidopsis thaliana] NCBI Description 2467 Seq. No. 1704 1.R1040 Contig ID

5'-most EST LIB3030-009-Q1-B1-H12

Method BLASTX
NCBI GI g1946367
BLAST score 390
E value 6.0e-38
Match length 91
% identity 80

NCBI Description (U93215) unknown protein [Arabidopsis thaliana]

Seq. No. 2468

Contig ID 1709_1.R1040 5'-most EST smc700745026.h1

Method BLASTX
NCBI GI g118514
BLAST score 2177
E value 0.0e+00
Match length 504
% identity 84

NCBI Description TURGOR-RESPONSIVE PROTEIN 26G >gi_100051_pir__S11863

aldehyde dehydrogenase homolog - garden pea

>gi_20681_emb_CAA38243_ (X54359) 508 aa peptide [Pisum

satīvum]

Seq. No. 2469

Contig ID 1709_2.R1040

5'-most EST LIB3170-039-Q1-J1-E12

Method BLASTX
NCBI GI g2244919
BLAST score 385
E value 8.0e-37
Match length 139
% identity 55

NCBI Description (Z97339) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2470

Contig ID 1709_3.R1040

5'-most EST jC-gmro02800026f06a1

Method BLASTX
NCBI GI g2980793
BLAST score 611
E value 3.0e-63
Match length 151
% identity 75

NCBI Description (AL022197) putative protein [Arabidopsis thaliana]

Seq. No. 2471

Contig ID 1709_4.R1040



```
LIB3028-020-Q1-B1-D3
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2129774
BLAST score
                  1095
E value
                  1.0e-120
Match length
                  292
% identity
                  xyloglucan endotransglycosylase-related protein XTR4 -
NCBI Description
                  Arabidopsis thaliana (fragment) >gi_1244754 (U43486)
                  xyloglucan endotransglycosylase-related protein
                   [Arabidopsis thaliana]
                  2472
Seq. No.
                  1709 5.R1040
Contig ID
5'-most EST
                  jex700905796.h1
Method
                  BLASTX
                  q3980400
NCBI GI
BLAST score
                  872
                   1.0e-93
E value
Match length
                   261
% identity
                   (AC004561) putative tropinone reductase [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                   2473
Contig ID
                   1709 6.R1040
                   jC-gmle01810084f07a1
5'-most EST
Method
                   BLASTN
                   q20680
NCBI GI
BLAST score
                   180
                   2.0e-96
E value
                   344
Match length
                   88
% identity
                  P.sativum mRNA of cDNA clone 26g
NCBI Description
                   2474
Seq. No.
                   1709_7.R1040
Contig ID
5'-most EST
                   dpv701099710.hl
Method
                   BLASTX
                   g2980793
NCBI GI
                   506
BLAST score
E value
                   3.0e-51
Match length
                   161
                   56
% identity
                  (AL022197) putative protein [Arabidopsis thaliana]
NCBI Description
                   2475
Seq. No.
                   1709 8.R1040
Contig ID
                   jC-gmle01810029a07d1
5'-most EST
```

BLASTX Method NCBI GI q2129774 BLAST score 214 E value 8.0e-17 Match length 75 % identity 52

xyloglucan endotransglycosylase-related protein XTR4 -NCBI Description Arabidopsis thaliana (fragment) >gi_1244754 (U43486)



xyloglucan endotransglycosylase-related protein
[Arabidopsis thaliana]

2476 Seq. No. 1709 10.R1040 Contig ID epx701109280.h1 5'-most EST Method BLASTX NCBI GI g2129774 BLAST score 324 4.0e-30 E value 92 Match length 66 % identity

NCBI Description xyloglucan endotransglycosylase-related protein XTR4 - Arabidopsis thaliana (fragment) >gi_1244754 (U43486)

xyloglucan endotransglycosylase-related protein

5

[Arabidopsis thaliana]

Seq. No. 2477

Contig ID 1709_11.R1040 5'-most EST rca701000554.h1 Method BLASTX

Method BLASTX
NCBI GI g2980793
BLAST score 322
E value 1.0e-29
Match length 118
% identity 58

NCBI Description (AL022197) putative protein [Arabidopsis thaliana]

Seq. No.

Contig ID 1709_14.R1040

5'-most EST LIB3106-003-Q1-K1-B1

2478

Method BLASTX
NCBI GI g2980793
BLAST score 213
E value 4.0e-17
Match length 106
% identity 44

NCBI Description (AL022197) putative protein [Arabidopsis thaliana]

Seq. No. 2479

Contig ID 1710_1.R1040 5'-most EST fua701043156.h1

Method BLASTX
NCBI GI g2497539
BLAST score 341
E value 2.0e-34
Match length 109
% identity 71

NCBI Description PYRUVATE KINASE, CHLOROPLAST ISOZYME A PRECURSOR >gi_169703

(M64736) ATP:pyruvate phosphotransferase [Ricinus communis]

Seq. No. 2480

Contig ID 1711_1.R1040 5'-most EST pmv700891332.h1

Method BLASTX NCBI GI g3559805 BLAST score 509



9.0e-52 E value Match length 142 % identity 63

(AJ006787) putative phytochelatin synthetase [Arabidopsis NCBI Description

thaliana]

2481 Seq. No.

1713 1.R1040 Contig ID $crh7\overline{0}0854175.h1$ 5'-most EST

BLASTX Method g3482971 NCBI GI BLAST score 533 6.0e-54E value 107 Match length % identity 87

(AL031369) putative protein [Arabidopsis thaliana] NCBI Description

2482 Seq. No.

1714 1.R1040 Contig ID

LIB3039-024-Q1-E1-G5 5'-most EST

Method BLASTX g2780194 NCBI GI 1713 BLAST score 0.0e + 00E value 388 Match length % identity 85

(AJ003197) adenine nucleotide translocator [Lupinus albus] NCBI Description

2483 Seq. No.

1714 2.R1040 Contig ID 5'-most EST ssr700557603.hl

2484 Seq. No.

1714_3.R1040 Contig ID

5'-most EST uC-gmflminsoy064c10b1

BLASTN Method NCBI GI q2780193 122 BLAST score 4.0e-62 E value Match length 378 % identity

NCBI Description Lupinus albus mRNA for adenine nucleotide translocator

2485 Seq. No.

1714_7.R1040 Contig ID kmv700741313.h15'-most EST

BLASTX Method NCBI GI g2780194 BLAST score 192 E value 1.0e-14 Match length 129 53 % identity

(AJ003197) adenine nucleotide translocator [Lupinus albus] NCBI Description

Seq. No.

2486 1715 1.R1040 Contig ID ncj700975104.h1 5'-most EST



Method BLASTX
NCBI GI g3236259
BLAST score 1249
E value 1.0e-138
Match length 289
% identity 80
NCBI Description (AC004684) putative alcohol dehydrogenase [Arabidopsis thaliana]

Seq. No. 2487

Contig ID

1717 1.R1040

5'-most EST

LIB3139-086-P1-N1-C11

Seq. No.

2488

Contig ID 5'-most EST 1717_2.R1040 jex700904932.h1

Method BLASTX
NCBI GI g3122139
BLAST score 327
E value 3.0e-30
Match length 137
% identity 48

NCBI Description

GLYCEROL KINASE (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)

(GLYCEROKINASE) (GK) (ATP-STIMULATED

GLUCOCORTICOID-RECEPTOR TRANSLOCATON PROMOTER) (ASTP)

>gi 484372 pir JN0606 ATP-stimulated

glucocorticoid-receptor translocation promoter protein - rat >gi_433211_dbj_BAA03677_ (D16102) ATP-stimulated glucocorticoid-receptor translocaton promoter [Rattus

rattus]

Seq. No.

2489

Contig ID 5'-most EST

1717_3.R1040 fde700874271.h1

Method BLASTX
NCBI GI g631174
BLAST score 291
E value 1.0e-26
Match length 78
% identity 69

NCBI Description

glycerol kinase - human >gi_515029_emb_CAA55364_ (X78711)

glycerol kinase [Homo sapiens]

Seq. No. 2490

Contig ID 5'-most EST 1717_4.R1040 eep700863731.h1

Seq. No. 2491

Contig ID

1719 1.R1040

5'-most EST

LIB3170-034-Q1-J1-C12

Seq. No.

2492

Contig ID

1719 2.R1040

5'-most EST

LIB3074-031-Q1-K1-F3

Seq. No.

2493

Contig ID

1719_3.R1040

NCBI Description



```
LIB3170-027-Q2-K1-B8
5'-most EST
Seq. No.
                  2494
Contig ID
                  1720 1.R1040
5'-most EST
                  rlr700900743.h1
                  BLASTX
Method
                  q3885343
NCBI GI
BLAST score
                  211
                  2.0e-16
E value
Match length
                  80
% identity
                  47
                  (AC005623) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  2495
Seq. No.
Contig ID
                  1720 2.R1040
5'-most EST
                  LIB3040-027-Q1-E1-E6
Method
                  BLASTX
                  q3885343
NCBI GI
BLAST score
                  229
                  7.0e-19
E value
Match length
                  119
% identity
                   43
                  (AC005623) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  2496
Seq. No.
                  1723_1.R1040
Contig ID
5'-most EST
                  LIB3109-017-Q1-K1-E9
                  BLASTX
Method
                   g498902
NCBI GI
                  586
BLAST score
                   1.0e-60
E value
                  135
Match length
                   84
% identity
                  (U10044) ribosomal protein L27 homolog [Pisum sativum]
NCBI Description
                   2497
Seq. No.
Contig ID
                   1723 2.R1040
                  LIB3073-026-Q1-K1-D11
5'-most EST
                   BLASTX
Method
                   q498902
NCBI GI
                   570
BLAST score
E value
                   1.0e-58
Match length
                   135
% identity
                   81
                  (U10044) ribosomal protein L27 homolog [Pisum sativum]
NCBI Description
                   2498
Seq. No.
                   1723 4.R1040
Contig ID
5'-most EST
                   zhf700959381.h1
                   BLASTX
Method
NCBI GI
                   q498902
BLAST score
                   150
E value
                   4.0e-13
Match length
                   46
% identity
                   89
```

(U10044) ribosomal protein L27 homolog [Pisum sativum]



2499 Seq. No. Contig ID

1724 1.R1040 leu701148868.h1

Method NCBI GI

5'-most EST

BLASTX q3334405

BLAST score E value

865 6.0e-93 239

Match length % identity

76

NCBI Description

VACUOLAR ATP SYNTHASE SUBUNIT E (V-ATPASE E SUBUNIT) >gi 2267583 (AF009338) vacuolar H+-ATPase subunit E

[Gossypium hirsutum]

Seq. No.

2500

Contig ID

1724 2.R1040

5'-most EST

jC-gmst02400048e07a1 BLASTX

Method NCBI GI BLAST score

g3334405 920

E value Match length

2.0e-99 232

81

% identity NCBI Description

VACUOLAR ATP SYNTHASE SUBUNIT E (V-ATPASE E SUBUNIT)

>qi 2267583 (AF009338) vacuolar H+-ATPase subunit E

[Gossypium hirsutum]

Seq. No.

2501

Contig ID

1724 4.R1040

5'-most EST

jC-gmle01810032d08d1

Method NCBI GI BLASTX q3334411

BLAST score E value

227 1.0e-18

Match length

51

% identity

NCBI Description

VACUOLAR ATP SYNTHASE SUBUNIT E (V-ATPASE E SUBUNIT) >qi 1263911 emb CAA65581 (X96785) vacuolar H(+)-ATPase

[Spinacia oleracea]

Seq. No.

2502

Contig ID

1724 5.R1040

5'-most EST

jC-gmro02800039d11d1

Seq. No.

2503

Contig ID

1724 7.R1040 asn701142409.hl

5'-most EST

BLASTX

Method NCBI GI

g3334405

BLAST score

471

E value

2.0e-47

Match length

107 89

% identity NCBI Description

VACUOLAR ATP SYNTHASE SUBUNIT E (V-ATPASE E SUBUNIT) >gi 2267583 (AF009338) vacuolar H+-ATPase subunit E

[Gossypium hirsutum]

Seq. No.

2504



```
1724 8.R1040
Contig ID
                  jC-gmro02910004a09d1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g1263910
BLAST score
                  72
                  6.0e-32
E value
Match length
                  152
% identity
                  87
                  S.oleracea mRNA for vacuolar H(+)-ATPase
NCBI Description
Seq. No.
                  2505
                  1725 1.R1040
Contig ID
                  LIB3028-054-Q1-B1-A2
5'-most EST
Seq. No.
                  2506
Contig ID
                  1727 1.R1040
                  LIB3030-001-Q1-B1-D8
5'-most EST
                   2507
Seq. No.
                  1728 1.R1040
Contig ID
5'-most EST
                  pmv700889401.h1
Method
                   BLASTX
                   g4538987
NCBI GI
BLAST score
                   538
                   1.0e-76
E value
Match length
                   286
% identity
                   58
                   (AJ133777) gamma-adaptin 2 [Arabidopsis thaliana]
NCBI Description
                   2508
Seq. No.
                   1730 1.R1040
Contig ID
                   q4300344
5'-most EST
                   BLASTX
Method
                   g1362093
NCBI GI
BLAST score
                   666
                   1.0e-69
E value
                   177
Match length
% identity
                   72
                  hypothetical protein (clone TPP15) - tomato (fragment)
NCBI Description
                   >qi 924632 (U20595) unknown [Solanum lycopersicum]
                   2509
Seq. No.
                   1732 1.R1040
Contig ID
                   LIB3107-044-Q1-K1-D12
5'-most EST
Method
                   BLASTX
                   q2739002
NCBI GI
                   1109
BLAST score
                   1.0e-121
E value
Match length
                   451
                   47
% identity
                  (AF022460) CYP83D1p [Glycine max]
NCBI Description
```

1735 1.R1040 Contig ID

5'-most EST LIB3039-034-Q1-E1-F4

2510

Method BLASTN NCBI GI g829118



```
BLAST score
                  1.0e-131
E value
                  550
Match length
                  87
% identity
NCBI Description P.vulgaris gene for cyclophilin
                  2511
Seq. No.
                  1735 2.R1040
Contig ID
5'-most EST
                  LIB3139-050-P1-N1-H3
                  BLASTN
Method
                  g829118
NCBI GI
BLAST score
                  303
                  1.0e-169
E value
Match length
                  631
% identity
                  88
NCBI Description P.vulgaris gene for cyclophilin
                  2512
Seq. No.
                  1735 3.R1040
Contig ID
5'-most EST
                   dpv701101486.hl
                  BLASTX
Method
                   g1703292
NCBI GI
                  357
BLAST score
                  2.0e-33
E value
Match length
                  187
                   70
% identity
NCBI Description HIGH AFFINITY AMMONIUM TRANSPORTER >gi_551219_emb_CAA53473_
                   (X75879) amt1 [Arabidopsis thaliana]
                   2513
Seq. No.
                   1735 5.R1040
Contig ID
5'-most EST
                   jC-qmro02910046f04d1
                   BLASTX
Method
NCBI GI
                   g4324714
                   938
BLAST score
E value
                   0.0e + 00
Match length
                   516
                   67
% identity
NCBI Description
                  (AF110771) ammonium transporter [Arabidopsis thaliana]
                   2514
Seq. No.
                   1735 7.R1040
Contig ID
5'-most EST
                   LIB3029-009-Q1-B1-G10
Method
                   BLASTX
NCBI GI
                   g1703292
BLAST score
                   762
E value
                   5.0e-81
Match length
                   225
% identity
                  HIGH AFFINITY AMMONIUM TRANSPORTER >gi 551219 emb CAA53473
NCBI Description
                   (X75879) amt1 [Arabidopsis thaliana]
                   2515
```

1735 8.R1040 Contig ID

5'-most EST jC-gmle01810000c04a1

BLASTX Method NCBI GI g1703292



```
BLAST score
E value
                  3.0e-90
Match length
                  230
                  72
% identity
NCBI Description
```

HIGH AFFINITY AMMONIUM TRANSPORTER >gi 551219 emb CAA53473 (X75879) amt1 [Arabidopsis thaliana]

Seq. No. 2516

1735_9.R1040 Contig ID 5'-most EST jC-gmle01810029c05d1

Method BLASTX g2392771 NCBI GI BLAST score 316 E value 4.0e-29 93 Match length 70 % identity

(AC002534) putative kinesin-like protein [Arabidopsis NCBI Description

thaliana]

Seq. No.

2517

Contig ID 1735_14.R1040

5'-most EST LIB3039-029-Q1-E1-B12

Method BLASTN NCBI GI g551218 BLAST score 80 E value 5.0e-37 Match length 160 88 % identity

NCBI Description A.thaliana amt1 mRNA

Seq. No.

2518

Contig ID 5'-most EST 1735 16.R1040 fua701036958.h1

Method BLASTX NCBI GI g4322319 BLAST score 194 E value 8.0e-30 Match length 89 % identity 74

(AF080541) ammonium transporter [Nepenthes alata] NCBI Description

2519 Seq. No.

1738 1.R1040 Contig ID 5'-most EST seb700648478.hl

Method BLASTX NCBI GI g4240207 BLAST score 264 E value 1.0e-22 172 Match length % identity 36

(AB020666) KIAA0859 protein [Homo sapiens] NCBI Description

Seq. No.

2520

2521

1738 2.R1040 Contig ID 5'-most EST gsv701056362.h1

Seq. No.

% identity

86



```
1739 1.R1040
 Contig ID
 5'-most EST
                    jex700903805.h1
 Method
                    BLASTX
 NCBI GI
                    g1857447
 BLAST score
                    338
 E value
                    2.0e-31
 Match length
                    162
 % identity
                    41
 NCBI Description
                    (U82367) UDP-glucose glucosyltransferase [Solanum
                    tuberosum]
                    2522
 Seq. No.
 Contig ID
                    1739 2.R1040
 5'-most EST
                    LIB3055-011-Q1-N1-A11
 Method
                    BLASTX
 NCBI GI
                    g1359905
 BLAST score
                  . 883
 E value
                    6.0e-95
 Match length
                    385
 % identity
                    46
 NCBI Description
                    (X85138) homologous to glucosyltransferases [Lycopersicon
                    esculentum]
 Seq. No.
                    2523
 Contig ID
                    1741 1.R1040
5'-most EST
                    q431\overline{3}526
 Method
                    BLASTN
                    g2225882
 NCBI GI
 BLAST score
                    183
 E value
                    4.0e-98
 Match length
                    483
                    84
 % identity
 NCBI Description
                    Solanum tuberosum mRNA for eukaryotic initiation factor
                    5A4, complete cds
 Seq. No.
                    2524
 Contig ID
                    1741 2.R1040
 5'-most EST
                    epx701106183.h1
 Method
                    BLASTN
 NCBI GI
                    q309672
 BLAST score
                    253
 E value
                    1.0e-140
 Match length
                    669
 % identity
 NCBI Description
                    Pisum sativum chloroplast photosystem I 24 kDa light
                    harvesting protein (lhca3) mRNA, complete cds
 Seq. No.
                    2525
                    1741 3.R1040
 Contig ID
 5'-most EST
                    LIB3051-055-Q1-K1-B5
 Method
                    BLASTN
 NCBI GI
                    g309672
 BLAST score
                    82
 E value
                    7.0e-38
 Match length
                    245
```

NCBI Description Pisum sativum chloroplast photosystem I 24 kDa light





harvesting protein (lhca3) mRNA, complete cds

Seq. No. 2526 1741 4.R1040 Contig ID 5'-most EST uC-gmflminsoy019f07b1 Method BLASTN NCBI GI g309672 BLAST score 98 E value 1.0e-47 Match length 255 87 % identity NCBI Description Pisum sativum chloroplast photosystem I 24 kDa light harvesting protein (lhca3) mRNA, complete cds 2527 Seq. No. Contig ID 1741 11.R1040 zhf700953944.h1 5'-most EST Method BLASTX NCBI GI q124224 BLAST score 528 E value 5.0e-54 Match length 102 % identity 96 NCBI Description INITIATION FACTOR 5A-1 (EIF-5A) (EIF-4D) >gi_100345_pir__S21060 translation initiation factor eIF-5A - common tobacco >gi_19887_emb_CAA45105_ (X63543) eukaryotic initiatin factor 5A (3) [Nicotiana tabacum] Seq. No. 2528 Contig ID 1742 1.R1040 5'-most EST pcp700994206.h1 Method BLASTX NCBI GI g4510406 BLAST score 1145 E value 1.0e-125 Match length 362 % identity (AC006587) putative protein kinase [Arabidopsis thaliana] NCBI Description 2529 Seq. No. 1742 2.R1040 Contig ID 5'-most EST uC-gmropic099d01b1 BLASTX Method NCBI GI g4580398 BLAST score 253 E value 1.0e-21 74 Match length % identity 68 NCBI Description (AC007171) putative protein kinase APK1A [Arabidopsis thaliana] 2530 Seq. No. 1742 3.R1040 Contig ID 5'-most EST smc700749001.h1

Method BLASTN NCBI GI q1209533 BLAST score 33



E value 5.0e-09 Match length 33 % identity 2

NCBI Description Mus musculus H2-M alpha chain (H2-Ma) gene, H2-M beta 2 chain (H2-Mb2) gene, H2-M beta 1 chain (H2-Mb1) gene, low molecular weight protein 2 Lmp2 (Lmp2) gene, complete cds

Seq. No. 2531

Contig ID 1746_1.R1040 5'-most EST V4L-01-Q1-B1-G3

Seq. No. 2532

Contig ID 1749 1.R1040

5'-most EST LIB3107-057-Q1-K1-F6

Method BLASTX
NCBI GI g2384956
BLAST score 151
E value 1.0e-17
Match length 183
% identity 38

NCBI Description (AF022985) No definition line found [Caenorhabditis

elegans]

Seq. No. 2533

Contig ID 1751 1.R1040

5'-most EST uC-gmflminsoy024e07b1

Method BLASTX
NCBI GI g1653622
BLAST score 379
E value 4.0e-36
Match length 167
% identity 50

NCBI Description (D90915) hypothetical protein [Synechocystis sp.]

Seq. No. 2534

Contig ID 1752_1.R1040

5'-most EST jC-gmle01810086c07a1

Method BLASTX
NCBI GI g3024908
BLAST score 222
E value 1.0e-36
Match length 193
% identity 47

NCBI Description HYPOTHETICAL 27.2 KD PROTEIN SLL0875

>gi 1652675 dbj BAA17595 (D90907) hypothetical protein

[Synechocystis sp.]

Seq. No. 2535

Contig ID 1752 2.R1040

5'-most EST LIB3028-053-Q1-B1-H3

Seq. No. 2536

Contig ID 1753_1.R1040 5'-most EST zhf700955476.h1

Method BLASTX NCBI GI g131528 BLAST score 765



E value 5.0e-81 Match length 442 % identity 16

NCBI Description POLYPYRIMIDINE TRACT-BINDING PROTEIN (PTB) (HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN I) (HNRNP I) (57 KD RNA-BINDING

PROTEIN PPTB-1) >gi_35768_emb_CAA43973_ (X62006) polypirimidine tract binding protein [Homo sapiens] >gi 35774_emb_CAA43056_ (X60648) polypyrimidine

tract-binding protein (pPTB) [Homo sapiens] >gi_4096061 (AC006273) PTB_HUMAN; PTB; HETEROGENEOUS NUCLEA; HNRNP I;

57 KD RNA-BINDING PROTEIN PPTB-1 [Homo sapiens]

Seq. No. 2537

Contig ID 1753_2.R1040

5'-most EST LIB3107-064-Q1-K1-G1

Method BLASTX
NCBI GI g91152
BLAST score 203
E value 1.0e-15
Match length 82
% identity 24

NCBI Description nuclear protein, 25K - mouse

Seq. No. 2538

Contig ID 1753 3.R1040

5'-most EST LIB3107-008-Q1-K1-F5

Seq. No. 2539

Contig ID 1754 2.R1040

5'-most EST LIB3109-046-Q1-K1-F4

Seq. No. 2540

Contig ID 1755 1.R1040

5'-most EST LIB3074-033-Q1-K1-A9

Method BLASTX
NCBI GI g3023436
BLAST score 956
E value 1.0e-103
Match length 243
% identity 73

NCBI Description CAFFEOYL-COA O-METHYLTRANSFERASE (TRANS-CAFFEOYL-COA

3-O-METHYLTRANSFERASE) (CCOAMT) (CCOAOMT) >gi_857578 (U27116) caffeoyl-CoA 3-O-methyltransferase [Populus

tremuloides]

Seq. No. 2541

Contig ID 1755_2.R1040 5'-most EST leu701155444.h1

Method BLASTX
NCBI GI g684942
BLAST score 1044
E value 1.0e-114
Match length 242
% identity 79

NCBI Description (U20736) S-adenosyl-L-methionine:trans-caffeoyl-CoA 3-O-methyltransferase [Medicago sativa subsp. sativa]

```
Seq. No.
                   2542
Contig ID
                  1755 3.R1040
                  LIB3107-005-Q1-K1-H11
5'-most EST
Method
                  BLASTX
NCBI GI
                  g533121
BLAST score
                  201
                  1.0e-15
E value
Match length
                  44
                  86
% identity
                  (U13151) S-adenosyl-L-methionine:trans-caffeoyl-CoA
NCBI Description
                  3-O-methyltransferase [Zinnia elegans]
                  2543
Seq. No.
Contig ID
                  1755 6.R1040
5'-most EST
                  g5753226
Method
                  BLASTN
                  g2995933
NCBI GI
BLAST score
                  34
E value
                  2.0e-09
Match length
                  58
                  90
% identity
                  Mesembryanthemum crystallinum caffeoyl-CoA
NCBI Description
                  O-methyltransferase mRNA, complete cds
Seq. No.
                  2544
Contig ID
                  1756 1.R1040
                  LIB3139-017-P1-N1-A11
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3860249
BLAST score
                   432
                   4.0e-42
E value
Match length
                  194
                   49
% identity
NCBI Description
                  (AC005824) unknown protein [Arabidopsis thaliana]
Seq. No.
                   2545
                   1756 3.R1040
Contig ID
5'-most EST
                  dpv701097848.h1
Seq. No.
                   2546
                   1757 1.R1040
Contiq ID
5'-most EST
                   uC-gmflminsoy040g12b1
Method
                  BLASTX
NCBI GI
                   g1652601
BLAST score
                   264
E value
                   2.0e-22
Match length
                   235
% identity
                   35
NCBI Description (D90906) hypothetical protein [Synechocystis sp.]
                   2547
Seq. No.
                   1757 2.R1040
Contig ID
```

5'-most EST

 $jsh7\overline{0}1068824.h1$

Seq. No.

2548

Contig ID 5'-most EST

1759 1.R1040 rca701001537.h1



```
Seq. No.
                  2549
                  1759 3.R1040
Contig ID
5'-most EST
                  LIB3073-007-Q1-K1-B3
Seq. No.
                  2550
Contig ID
                  1760 1.R1040
5'-most EST
                  LIB3049-030-Q1-E1-E4
Method
                  BLASTX
NCBI GI
                  q2983997
BLAST score
                  426
E value
                  1.0e-41
Match length
                  159
% identity
                  52
                  (AE000749) hypothetical protein [Aquifex aeolicus]
NCBI Description
                  2551
Seq. No.
Contig ID
                  1762 1.R1040
5'-most EST
                  smc7\overline{0}0748901.h1
Method
                  BLASTX
NCBI GI
                  g3176664
BLAST score
                  510
E value
                   2.0e-51
Match length
                  153
% identity
                   60
                   (AC004393) Contains similarity to beta scruin gb Z47541
NCBI Description
                   from Limulus polyphemus. ESTs gb T04493 and gb_AA585955
                  come from this gene. [Arabidopsis thaliana]
                   2552
Seq. No.
                   1764 1.R1040
Contig ID
5'-most EST
                  txt700733348.hl
Seq. No.
                   2553
Contig ID
                   1765 1.R1040
                   zhf700960918.h1
5'-most EST
Method
                   BLASTX
                   g1652856
NCBI GI
BLAST score
                   300
                   8.0e-27
E value
Match length
                   162
% identity
                   41
                  (D90909) hypothetical protein [Synechocystis sp.]
NCBI Description
                   2554
Seq. No.
                   1766 1.R1040
Contig ID
5'-most EST
                  LIB3028-053-Q1-B1-E7
Method
                  BLASTX
NCBI GI
                   q4006896
BLAST score
                   162
                   2.0e-15
E value
Match length
                   76
                   64
% identity
```

Seq. No. 2555

NCBI Description

Contig ID 1767 1.R1040

(Z99708) SCARECROW-like protein [Arabidopsis thaliana]



```
epx701108250.h1
5'-most EST
Method
                   BLASTX
                   q4158221
NCBI GI
BLAST score
                   1715
                   0.0e + 00
E value
                   354
Match length
                   86
% identity
                   (Y18624) reversibly glycosylated polypeptide [Oryza sativa]
NCBI Description
                   2556
Seq. No.
                   1768 1.R1040
Contig ID
                   jC-gmro02910054c10a1
5'-most EST
                   BLASTX
Method
                   q3763933
NCBI GI
BLAST score
                   283
                   5.0e-25
E value
Match length
                   67
                   84
% identity
                  (AC004450) unknown protein [Arabidopsis thaliana]
NCBI Description
                   2557
Seq. No.
                   1769 1.R1040
Contig ID
                   ncj700983042.h1
5'-most EST
Method
                   BLASTN
NCBI GI
                   g2827081
                   601
BLAST score
                   0.0e+00
E value
                   1016
Match length
% identity
                   91
                  Medicago sativa cytosolic malate dehydrogenase (cmdh) mRNA,
NCBI Description
                   complete cds
                   2558
Seq. No.
Contig ID
                   1769 2.R1040
                   fua701043420.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4115534
                   1466
BLAST score
                   1.0e-176
E value
Match length
                   400
% identity
                   75
                   (AB012114) UDP-glycose:flavonoid glycosyltransferase [Vigna
NCBI Description
                   mungo]
Seq. No.
                   2559
                   1769 3.R1040
Contig ID
5'-most EST
                   q440\overline{5}695
Method
                   BLASTN
NCBI GI
                   g2827081
BLAST score
                   94
E value
                   4.0e-45
```

F ... 1.00

Match length 303 % identity

Medicago sativa cytosolic malate dehydrogenase (cmdh) mRNA, NCBI Description

complete cds

2560 Seq. No.



```
1770 1.R1040
Contig ID
                  fua701036901.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1209020
BLAST score
                  285
E value
                  7.0e-25
                  223
Match length
                  34
% identity
                  (U09584) PL6 protein [Homo sapiens]
NCBI Description
Seq. No.
                  2561
                  1770 2.R1040
Contig ID
                  leu701144340.h1
5'-most EST
                  2562
Seq. No.
Contig ID
                  1771 1.R1040
                  LIB3167-050-P1-K1-C12
5'-most EST
Method
                  BLASTX
                  g1363479
NCBI GI
BLAST score
                  742
                  1.0e-78
E value
Match length
                   205
                   72
% identity
                  photosystem I protein psaL - cucumber
NCBI Description
                  >gi 801740 dbj_BAA09047_ (D50456) PsaL [Cucumis sativus]
                   2563
Seq. No.
                   1771 3.R1040
Contig ID
                   LIB3138-087-P1-N1-A8
5'-most EST
                  BLASTX
Method
                   q1363479
NCBI GI
BLAST score
                   254
                   3.0e-25
E value
                   100
Match length
                   63
% identity
                   photosystem I protein psaL - cucumber
NCBI Description
                   >gi_801740_dbj_BAA09047_ (D50456) PsaL [Cucumis sativus]
                   2564
Seq. No.
                   1771 4.R1040
Contig ID
                   zhf700960439.h1
5'-most EST
                   BLASTX
Method
                   g1363479
NCBI GI
                   266
BLAST score
                   3.0e-23
E value
Match length
                   74
                   76
% identity
                   photosystem I protein psaL - cucumber
NCBI Description
                   >gi 801740_dbj_BAA09047_ (D50456) PsaL [Cucumis sativus]
Seq. No.
                   2565
                   1771 5.R1040
Contig ID
5'-most EST
                   vzy700752230.h1
                   2566
Seq. No.
```

1773 1.R1040

jC-gmro02910029d07a1

Contig ID

5'-most EST



```
BLASTX
Method
                   q4239845
NCBI GI
BLAST score
                   2129
E value
                   0.0e + 00
Match length
                   611
% identity
                   68
NCBI Description
```

(AB015855) transcription factor TEIL [Nicotiana tabacum]

Seq. No. 2567

1773 2.R1040 Contig ID

 $LIB3\overline{0}28-002-Q1-B1-F6$ 5'-most EST

Method BLASTX g4239845 NCBI GI BLAST score 611 E value 2.0e-63 Match length 144 % identity 84

(AB015855) transcription factor TEIL [Nicotiana tabacum] NCBI Description

2568 Seq. No.

Contig ID 1773 3.R1040

uC-qmropic019g06b1 5'-most EST

Method BLASTX NCBI GI g4239845 BLAST score 200 E value 2.0e-15 Match length 58 % identity 67

(AB015855) transcription factor TEIL [Nicotiana tabacum] NCBI Description

Seq. No.

2569 1773 4.R1040 Contig ID zhf700964078.h1 5'-most EST

BLASTX Method g4239845 NCBI GI 290 BLAST score 7.0e-26 E value 80 Match length 74 % identity

(AB015855) transcription factor TEIL [Nicotiana tabacum] NCBI Description

2570 Seq. No.

1773 5.R1040 Contig ID

LIB3093-058-Q1-K1-E4 5'-most EST

BLASTX Method NCBI GI g4239845 BLAST score 205 E value 4.0e-16Match length 59 68 % identity

(AB015855) transcription factor TEIL [Nicotiana tabacum] NCBI Description

Seq. No.

1773 6.R1040 Contig ID 5'-most EST hyd700726505.h1

2571

Method BLASTX NCBI GI g4239845



BLAST score 254
E value 1.0e-21
Match length 101
% identity 49

NCBI Description (AB015855) transcription factor TEIL [Nicotiana tabacum]

Seq. No. 2572

Contig ID 1773 7.R1040

5'-most EST LIB3051-037-Q1-K1-H5

Method BLASTX
NCBI GI g4239845
BLAST score 554
E value 5.0e-57
Match length 144
% identity 75

NCBI Description (AB015855) transcription factor TEIL [Nicotiana tabacum]

Seq. No. 2573

Contig ID 1776 1.R1040

5'-most EST jC-gmst02400060h05a1

Seq. No. 2574

Contig ID 1776 2.R1040

5'-most EST jC-gmst02400051g07a1

Method BLASTX
NCBI GI g2967452
BLAST score 1734
E value 0.0e+00
Match length 701
% identity 56

NCBI Description (AB010882) hSNF2H [Homo sapiens]

>gi_4507075_ref_NP_003592.1_pSMARCA5_ SWI/SNF related,
matrix associated, actin dependent regulator of chromatin,

subfamily a, member

Seq. No. 2575

Contig ID 1776_3.R1040 5'-most EST pxt700942642.h1

Method BLASTX
NCBI GI g2735256
BLAST score 378
E value 3.0e-36
Match length 74
% identity 95

NCBI Description (U89683) protein kinase [Lycopersicon esculentum]

Seq. No. 2576

Contig ID 1778 1.R1040

5'-most EST LIB3170-046-Q1-J1-D6

Method BLASTX
NCBI GI g543905
BLAST score 1023
E value 1.0e-111
Match length 241
% identity 77

NCBI Description BRASSINOSTEROID-REGULATED PROTEIN BRU1 >gi_347459 (L22162)

brassinosteroid-regulated protein [Glycine max]



```
2577
Seq. No.
                   1778 2.R1040
Contig ID
                   uC-gmflminsoy078d11b1
5'-most EST
Method
                   BLASTX
                   q543905
NCBI GI
BLAST score
                   715
                   1.0e-75
E value
Match length
                   164
                   79
% identity
                   BRASSINOSTEROID-REGULATED PROTEIN BRU1 >gi_347459 (L22162)
NCBI Description
                   brassinosteroid-regulated protein [Glycine max]
                   2578
Seq. No.
Contig ID
                   1778 3.R1040
5'-most EST
                   g5666939
                   BLASTX
Method
                   g2129771
NCBI GI
BLAST score
                   351
                   4.0e-33
E value
                   90
Match length
                   76
% identity
                   xyloglucan endotransglycosylase-related protein XTR-6 -
NCBI Description
                   Arabidopsis thaliana >gi 1244758 (U43488) xyloglucan
                   endotransglycosylase-related protein [Arabidopsis thaliana]
                   >gi_4539299_emb_CAB39602.1_ (AL049480) xyloglucan endo-1,
4-beta-D-glucanase (XTR-6) [Arabidopsis thaliana]
                   2579
Seq. No.
Contig ID
                   1779 1.R1040
                   asn701141656.h1
5'-most EST
Seq. No.
                   2580
Contig ID
                   1780 1.R1040
5'-most EST
                   vzy700751721.h1
Method
                   BLASTX
NCBI GI
                   q4091080
BLAST score
                   520
E value
                   8.0e-53
Match length
                   141
% identity
                   (AF045571) nucleic acid binding protein [Oryza sativa]
NCBI Description
Seq. No.
                   2581
                   1784 1.R1040
Contig ID
5'-most EST
                   g4276976
Method
                   BLASTX
NCBI GI
                   g544134
BLAST score
                    355
```

E value 2.0e-33 Match length 183

% identity 22 NCBI Description

DNA-DAMAGE-REPAIR/TOLERATION PROTEIN DRT100 PRECURSOR >gi 99720 pir S22863 hypothetical protein - Arabidopsis thaliana >gi 421844 pir A46260 RecA functional analog

DRT100 - Arabidopsis thaliana (fragment)



 Seq. No.
 2582

 Contig ID
 1785_1.R1040

 5'-most EST
 LIB3109-036-Q1-K1-B11

Method BLASTX
NCBI GI g4206765
BLAST score 728
E value 9.0e-77
Match length 367

% identity 46
NCBI Description (AF104329) putative type 1 membrane protein [Arabidopsis

thaliana]

Seq. No. 2583

Contig ID 1785_3.R1040

5'-most EST LIB3051-112-Q1-K1-B8

Method BLASTX
NCBI GI g4206765
BLAST score 194
E value 5.0e-15
Match length 86
% identity 57

NCBI Description (AF104329) putative type 1 membrane protein [Arabidopsis

thaliana]

Seq. No. 2584

Contig ID 1787_1.R1040

5'-most EST jC-gmro02910040b01a1

Method BLASTX
NCBI GI g2245012
BLAST score 499
E value 1.0e-85
Match length 257
% identity 66

NCBI Description (Z97341) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2585

Contig ID 1789_1.R1040 5'-most EST pcp700991322.h1

Method BLASTX
NCBI GI g1903034
BLAST score 1274
E value 1.0e-140
Match length 544
% identity 48

NCBI Description (X94625) amp-binding protein [Brassica napus]

Seq. No. 2586

Contig ID 1790 1.R1040

5'-most EST LIB3028-053-Q1-B1-C5

Method BLASTX
NCBI GI g2244956
BLAST score 277
E value 2.0e-24
Match length 112
% identity 46

NCBI Description (Z97340) strong similarity to pectinesterase [Arabidopsis

thaliana]



Seq. No. 2587

Contig ID 1791 1.R1040

5'-most EST uC-gmrominsoy258e11b1

Method BLASTX
NCBI GI g2146732
BLAST score 2059
E value 0.0e+00
Match length 541
% identity 74

NCBI Description FK506-binding protein - Arabidopsis thaliana >gi_1373396

(U57838) rof1 [Arabidopsis thaliana]

Seq. No. 2588

Contig ID 1791 2.R1040

5'-most EST jC-qmst02400069c09a1

Method BLASTN
NCBI GI g1354206
BLAST score 76
E value 2.0e-34
Match length 212
% identity 84

NCBI Description Arabidopsis thaliana FK506 binding protein FKBP62 (ROF1)

mRNA, complete cds

Seq. No. 2589

Contig ID 1791_4.R1040 5'-most EST asn701140476.h1

Method BLASTX
NCBI GI g3023751
BLAST score 483
E value 1.0e-48
Match length 134
% identity 29

NCBI Description

cription 70 KD PEPTIDYLPROLYL ISOMERASE (PEPTIDYLPROLYL CIS-TRANS ISOMERASE) (CYCLOPHILIN) (PPIASE) >gi_1076772_pir__S55383

peptidylprolyl isomerase (EC 5.2.1.8) - wheat

>gi_854626_emb_CAA60505_ (X86903) peptidylprolyl isomerase

[Triticum aestivum]

Seq. No. 2590

Contig ID 1791_5.R1040 5'-most EST seb700650371.h1 Method BLASTN

NCBI GI g1354206 BLAST score 71 E value 2.0e-31 Match length 247 % identity 82

NCBI Description Arabidopsis thaliana FK506 binding protein FKBP62 (ROF1)

mRNA, complete cds

Seq. No. 2591

Contig ID 1791 7.R1040

5'-most EST LIB3138-007-Q1-N1-D6

Method BLASTX NCBI GI g2146731



BLAST score 487 E value 3.0e-49 Match length 124 % identity 29

NCBI Description FK506-binding protein - Arabidopsis thaliana >gi_1354207

(U49453) rof1 [Arabidopsis thaliana]

Seq. No. 2592

Contig ID 1791_8.R1040 5'-most EST leu701150772.h1

Method BLASTX
NCBI GI g3023751
BLAST score 207
E value 1.0e-16
Match length 63
% identity 47

NCBI Description 70 KD PEPTIDYLPROLYL ISOMERASE (PEPTIDYLPROLYL CIS-TRANS

ISOMERASE) (CYCLOPHILIN) (PPIASE) >gi_1076772_pir__S55383

peptidylprolyl isomerase (EC 5.2.1.8) - wheat

>gi_854626_emb_CAA60505_ (X86903) peptidylprolyl isomerase

[Triticum aestivum]

Seq. No. 2593

1793 1.R1040 Contig ID g5510050 5'-most EST Method BLASTX q3914685 NCBI GI BLAST score 614 E value 1.0e-63 143 Match length 83 % identity

NCBI Description 60S RIBOSOMAL PROTEIN L17 >gi_2668748 (AF034948) ribosomal

protein L17 [Zea mays]

Seq. No. 2594

Contig ID 1793 2.R1040

5'-most EST LIB3106-075-Q1-K2-B7

Method BLASTX
NCBI GI g3176668
BLAST score 743
E value 1.0e-78
Match length 174
% identity 82

NCBI Description (AC004393) Similar to ribosomal protein L17 gb_X62724 from

Hordeum vulgare. ESTs gb_Z34728, gb_F19974, gb_T75677 and

gb Z33937 come from this gene. [Arabidopsis thaliana]

Seq. No. 2595

Contig ID 1793_3.R1040

5'-most EST LIB3170-030-Q1-K1-H1

Method BLASTX
NCBI GI g131167
BLAST score 321
E value 2.0e-29
Match length 131
% identity 58

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT II PRECURSOR



(PHOTOSYSTEM I 20 KD SUBUNIT) (PSI-D)
>gi_100292_pir__S18348 photosystem I chain II precursor wood tobacco >gi_19748_emb_CAA42623_ (X60008) PSI-D2
[Nicotiana sylvestris]

Seq. No. 2596

Contig ID 1793 4.R1040

5'-most EST LIB3028-004-Q1-B1-D11

Method BLASTN
NCBI GI g19747
BLAST score 131
E value 5.0e-67
Match length 441
% identity 84

NCBI Description N.sylvestris psaDa gene for PSI-D2

Seq. No. 2597

Contig ID 1793 5.R1040

5'-most EST LIB3170-002-Q1-K1-A4

Method BLASTN
NCBI GI g1806284
BLAST score 193
E value 1.0e-104
Match length 313
% identity 90

NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2598

Contig ID 1793_6.R1040

5'-most EST LIB3093-038-Q1-K1-E10

Method BLASTX
NCBI GI g487046
BLAST score 770
E value 5.0e-82
Match length 208
% identity 75

NCBI Description photosystem I chain II precursor - wood tobacco >gi 407769 dbj BAA02871 (D13718) PSI-D1 precursor

[Nicotiana sylvestris]

Seq. No.

Contig ID 1793 7.R1040

5'-most EST LIB3051-095-Q1-K1-E8

2599

Method BLASTX
NCBI GI 9730526
BLAST score 502
E value 6.0e-90
Match length 192
% identity 84

NCBI Description 60S RIBOSOMAL PROTEIN L13 (BBC1 PROTEIN HOMOLOG)

>gi_480787_pir__S37271 ribosomal protein L13 - Arabidopsis
thaliana >gi_404166_emb_CAA53005_ (X75162) BBC1 protein

[Arabidopsis thaliana]

Seq. No.

2600

Contig ID 1793_8.R1040

5'-most EST LIB3170-013-Q1-K1-E11



Method BLASTN
NCBI GI g2351070
BLAST score 115
E value 1.0e-57
Match length 303
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTH12, complete sequence [Arabidopsis thaliana]

Seq. No. 2601

Contig ID 1793 9.R1040

5'-most EST LIB3092-006-Q1-K1-H1

Method BLASTX
NCBI GI g131166
BLAST score 574
E value 4.0e-59
Match length 172
% identity 69

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT II PRECURSOR

(PHOTOSYSTEM I 20 KD SUBUNIT) (PSI-D) >gi 82100 pir__S00449

photosystem I chain II precursor - tomato >gi_170492 (M21344) photosystem I subunit II protein precursor [Lycopersicon esculentum] >gi_226544_prf__1601516A

photosystem I reaction center II [Lycopersicon esculentum]

Seq. No. 2602

Contig ID 1793_10.R1040

5'-most EST LIB3049-007-Q1-E1-B3

Method BLASTN
NCBI GI g166739
BLAST score 97
E value 6.0e-47
Match length 209
% identity 87

NCBI Description A.thaliana histone H4 gene, complete cds

Seq. No. 2603

Contig ID 1793_11.R1040

5'-most EST LIB3167-078-P1-K2-H6

Method BLASTX
NCBI GI g487046
BLAST score 433
E value 1.0e-46
Match length 117
% identity 85

NCBI Description photosystem I chain II precursor - wood tobacco

>gi_407769_dbj_BAA02871_ (D13718) PSI-D1 precursor

[Nicotiana sylvestris]

Seq. No. 2604

Contig ID 1793 12.R1040

5'-most EST LIB3040-042-Q1-E1-G10

Method BLASTN
NCBI GI g1806284
BLAST score 185
E value 2.0e-99
Match length 313



% identity 90
NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2605

Contig ID 1793 13.R1040

5'-most EST LIB3106-049-Q1-K1-G7

Method BLASTX
NCBI GI g3914685
BLAST score 278
E value 2.0e-24
Match length 73
% identity 73

NCBI Description 60S RIBOSOMAL PROTEIN L17 >gi_2668748 (AF034948) ribosomal

protein L17 [Zea mays]

Seq. No. 2606

Contig ID 1793_14.R1040 5'-most EST pmv700892448.h1

Seq. No. 2607

Contig ID 1793_15.R1040

5'-most EST LIB3170-021-Q1-K1-C1

Method BLASTN
NCBI GI g2351070
BLAST score 125
E value 1.0e-63
Match length 305
% identity 85

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTH12, complete sequence [Arabidopsis thaliana]

Seq. No. 2608

Contig ID 1793_16.R1040

5'-most EST LIB3106-097-Q1-K1-G9

Method BLASTN
NCBI GI g1806284
BLAST score 144
E value 3.0e-75
Match length 312
% identity 87

NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2609

Contig ID 1793_17.R1040

5'-most EST LIB3049-032-Q1-E1-D8

Method BLASTN
NCBI GI g1806284
BLAST score 51
E value 8.0e-20
Match length 95
% identity 88

NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2610

Contig ID 1793_18.R1040

5'-most EST LIB3170-026-Q1-K1-B9

Method BLASTN

```
NCBI GI g166739
BLAST score 93
E value 9.0e-45
```

E value 9.0e-4
Match length 209
% identity 86

NCBI Description A.thaliana histone H4 gene, complete cds

Seq. No.

Contig ID 1793_19.R1040 5'-most EST jsh701068445.h1

Seq. No. 2612

Contig ID 1793 22.R1040

5'-most EST LIB3170-020-Q1-K1-H3

85

2611

Method BLASTN
NCBI GI g1806284
BLAST score 83
E value 8.0e-39
Match length 203

NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2613

% identity

Contig ID 1793_27.R1040 5'-most EST gsv701051886.h1

Method BLASTN
NCBI GI g1806284
BLAST score 138
E value 9.0e-72
Match length 234
% identity 90

NCBI Description S.rostrata mRNA coding for histone H4 homologue

Seq. No. 2614

Contig ID 1793_30.R1040 5'-most EST vwf700676687.h1

Method BLASTX
NCBI GI g730450
BLAST score 249
E value 1.0e-21
Match length 65
% identity 75

NCBI Description 60S RIBOSOMAL PROTEIN L13-2 (COLD INDUCED PROTEIN C24B)

>gi_480649_pir__S37134 cold-induced protein BnC24B - rape >gi_398922_emb_CAA80343_ (Z22620) cold induced protein

(BnC24B) [Brassica napus]

Seq. No. 2615

Contig ID 1793_33.R1040 5'-most EST uC-gmropic004h06b1

Method BLASTX
NCBI GI g3935177
BLAST score 155
E value 8.0e-11
Match length 32
% identity 91

NCBI Description (AC004557) F17L21.20 [Arabidopsis thaliana]



2616 Seq. No.

Contig ID 1794 1.R1040 5'-most EST pmv700888607.hl

Method BLASTN NCBI GI q344003 BLAST score 319 E value 1.0e-179 Match length 944 % identity 84

Pisum sativum mRNA for precursor for 33-kDa protein of NCBI Description

photosystem II, complete cds

2617 Seq. No.

Contig ID 1794 2.R1040 5'-most EST q4396866 Method BLASTX NCBI GI q131384 BLAST score 888 E value 9.0e-96 Match length 206 % identity 87

OXYGEN-EVOLVING ENHANCER PROTEIN 1 PRECURSOR (OEE1) (33 KD NCBI Description

SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II) (33 KD

THYLAKOID MEMBRANE PROTEIN) >gi 81934 pir S04132

photosystem II oxygen-evolving complex protein 1 precursor - garden pea >gi 20621 emb CAA33408 (X15350) precursor (AA

-81 to 248) [Pisum satīvum] >gi_344004_dbj_BAA02554_ (D13297) precursor for 33-kDa protein of photosystem II [Pisum sativum] >gi_226937_prf__1611461A O2 evolving

complex 33kD protein [Arachis hypogaea]

2618 Seq. No.

1794 3.R1040 Contig ID 5'-most EST $fde7\overline{0}0876743.h1$

Method BLASTX g131384 NCBI GI BLAST score 682 E value 8.0e-72 165 Match length % identity 84

OXYGEN-EVOLVING ENHANCER PROTEIN 1 PRECURSOR (OEE1) (33 KD NCBI Description

SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II) (33 KD

THYLAKOID MEMBRANE PROTEIN) >gi 81934_pir__S04132

photosystem II oxygen-evolving complex protein 1 precursor

- garden pea >gi_20621_emb_CAA33408_ (X15350) precursor (AA -81 to 248) [Pisum sativum] >gi_344004_dbj_BAA02554_ (D13297) precursor for 33-kDa protein of photosystem II [Pisum sativum] >gi 226937 prf 1611461A O2 evolving

complex 33kD protein [Arachis hypogaea]

2619 Seq. No.

1794 4.R1040 Contig ID

5'-most EST LIB3170-029-Q1-J1-A9

Method BLASTN NCBI GI g20620 BLAST score 134



E value 4.0e-69
Match length 286
% identity 87

NCBI Description Pea mRNA for the 33kDa polypeptide of the water-oxidizing

complex of photosystem II

Seq. No. 2620

Contig ID 1794_5.R1040

5'-most EST jC-gmle01810047h09d1

Method BLASTN
NCBI GI g344003
BLAST score 119
E value 3.0e-60
Match length 271
% identity 86

NCBI Description Pisum sativum mRNA for precursor for 33-kDa protein of

photosystem II, complete cds

Seq. No. 2621

Contig ID 1795 1.R1040 5'-most EST all700863141.h1

Method BLASTX
NCBI GI g3608152
BLAST score 202
E value 5.0e-16
Match length 78
% identity 54

NCBI Description (AC005314) hypothetical protein [Arabidopsis thaliana]

Seq. No. 2622

Contig ID 1798_1.R1040
5'-most EST g4299089

Method BLASTX

NCBI GI g3152585

BLAST score 323

E value 1.0e-29

Match length 140

Match length 140 % identity 55

NCBI Description (AC002986) Contains similarity to auxin-induced protein TM018A10.6 from A. thaliana BAC gb AF013294. [Arabidopsis

thaliana]

Seq. No. 2623

Contig ID 1798 2.R1040

5'-most EST uC-gmrominsoy216f03b1

Method BLASTX
NCBI GI g3152585
BLAST score 278
E value 3.0e-24
Match length 141
% identity 48

NCBI Description (AC002986) Contains similarity to auxin-induced protein

TM018A10.6 from A. thaliana BAC gb_AF013294. [Arabidopsis

thaliana]

Seq. No. 2624

Contig ID 1799_1.R1040



```
zzp700835068.hl
5'-most EST
                  BLASTX
Method
                  g4510376
NCBI GI
                   1206
BLAST score
                   1.0e-133
E value
Match length
                   414
% identity
                   57
                  (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1799 2.R1040
Contig ID
                   uC-gmropic050d06b1
5'-most EST
                   BLASTX
Method
                   g4510376
NCBI GI
BLAST score
                   414
                   2.0e-40
E value
Match length
                   106
% identity
NCBI Description (AC007017) unknown protein [Arabidopsis thaliana]
                   2626
Seq. No.
                   1800 1.R1040
Contig ID
                   leu701147459.h1
5'-most EST
                   BLASTX
Method
                   g2414622
NCBI GI
                   585
BLAST score
E value
                   2.0e-60
                   169
Match length
% identity
                   (Z99259) conserved protein [Schizosaccharomyces pombe]
NCBI Description
                   2627
Seq. No.
                   1800 2.R1040
Contig ID
                   ncj700987772.hl
5'-most EST
                   2628
Seq. No.
Contig ID
                   1802 1.R1040
                   zpv700761893.hl
5'-most EST
                   BLASTX
Method
                   g3201627
NCBI GI
BLAST score
                   177
E value
                   1.0e-12
Match length
                   33
                   91
% identity
                   (AC004669) putative SWH1 protein [Arabidopsis thaliana]
NCBI Description
                   2629
Seq. No.
                   1804_1.R1040
Contig ID
                   LIB3051-112-Q1-K1-A7
5'-most EST
                   BLASTX
Method
                   g70644
NCBI GI
BLAST score
                   1895
                   0.0e + 00
E value
Match length
                   380
                   53
% identity
```

NCBI Description ubiquitin precursor - common sunflower (fragment)

```
Seq. No.
                  1804 2.R1040
Contig ID
                  LIB3028-015-Q1-B1-E9
5'-most EST
                  BLASTN
Method
                  g516853
NCBI GI
                  726
BLAST score
E value
                   0.0e + 00
                   776
Match length
                   75
% identity
                  Soybean SUBI-2 gene for ubiquitin, complete cds
NCBI Description
                   2631
Seq. No.
                   1804 3.R1040
Contig ID
```

LIB3093-022-Q1-K1-G1 5'-most EST

BLASTN Method g516853 NCBI GI 888 BLAST score 0.0e+00E value 1070 Match length 49 % identity

Soybean SUBI-2 gene for ubiquitin, complete cds NCBI Description

Seq. No.

1804 4.R1040 Contig ID

LIB3050-016-Q1-E1-D6 5'-most EST

2632

BLASTN Method g1684856 NCBI GI BLAST score 199 1.0e-108 E value 299 Match length

38 % identity

NCBI Description Phaseolus vulgaris polyubiquitin mRNA, partial cds

2633 Seq. No.

1804 5.R1040 Contig ID pxt700945034.h1 5'-most EST

BLASTN Method NCBI GI q516853 BLAST score 215 1.0e-117 E value Match length 383 % identity 42

NCBI Description Soybean SUBI-2 gene for ubiquitin, complete cds

Seq. No. 2634

1804 7.R1040 Contig ID

LIB3051-053-Q1-K2-H9 5'-most EST

Method BLASTN g516853 NCBI GI BLAST score 180 9.0e-97 E value Match length 269 % identity 30

Soybean SUBI-2 gene for ubiquitin, complete cds NCBI Description

2635 Seq. No.

1805 1.R1040 Contig ID



```
uC-qmrominsoy048a03b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3123295
                   339
BLAST score
                   2.0e-31
E value
                   136
Match length
                   52
% identity
                   CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi 2583169
NCBI Description
                   (AF026473) calmodulin-related protein [Arabidopsis
                   thaliana]
                   2636
Seq. No.
Contig ID
                   1805 2.R1040
                   LIB3028-042-Q1-B1-H10
5'-most EST
                   BLASTX
Method
                   g3123295
NCBI GI
                   198
BLAST score
                   3.0e-15
E value
Match length
                   64
% identity
                   61
                   CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi 2583169
NCBI Description
                   (AF026473) calmodulin-related protein [Arabidopsis
                   thaliana]
Seq. No.
                   2637
                   1810 1.R1040
Contig ID
                   zhf7\overline{0}0955314.h1
5'-most EST
                   2638
Seq. No.
                   1811 1.R1040
Contig ID
                   g4291898
5'-most EST
                   2639
Seq. No.
                   1811 2.R1040
Contig ID
                   jC-qmro02910023d08d1
5'-most EST
                   2640
Seq. No.
                   1812 1.R1040
Contig ID
5'-most EST
                   LIB3028-053-Q1-B1-A5
                   2641
Seq. No.
                   1814 1.R1040
Contig ID
5'-most EST
                    zhf700955022.h1
                    2642
Seq. No.
Contig ID
                    1815 1.R1040
5'-most EST
                   uC-gmrominsoy275h01b1
Seq. No.
                    2643
                    1815 2.R1040
Contig ID
```

5'-most EST

uC-gmrominsoy157d05b1

Seq. No. Contig ID 2644

2645

1815 3.R1040

5'-most EST

LIB3092-006-Q1-K1-C12

Seq. No.

Seq. No.



```
1815 6.R1040
Contig ID
                  wrg700788505.hl
5'-most EST
                   2646
Seq. No.
                   1817 1.R1040
Contig ID
                  LIB3028-052-Q1-B2-D9
5'-most EST
Seq. No.
                   2647
                   1818 1.R1040
Contig ID
                   jC-qmle01810014a07d1
5'-most EST
                   2648
Seq. No.
                   1819 1.R1040
Contig ID
                   dpv701097046.h1
5'-most EST
```

BLASTX Method g3540183 NCBI GI 1022 BLAST score 1.0e-111 E value 304 Match length 65 % identity

(AC004122) Highly Similar to branched-chain amino acid NCBI Description aminotransferase [Arabidopsis thaliana]

1819 2.R1040 Contig ID uC-gmropic061b03b1 5'-most EST BLASTX Method g3540185 NCBI GI 186 BLAST score 8.0e-14 E value 88 Match length 47 % identity

2649

(AC004122) Highly Similar to branched-chain amino acid NCBI Description aminotransferase [Arabidopsis thaliana]

2650 Seq. No. 1819 3.R1040 Contig ID 5'-most EST xpa700794532.hl BLASTX Method q2827709 NCBI GI BLAST score 169

E value 6.0e-12 Match length 45 % identity

(AL021684) predicted protein [Arabidopsis thaliana] NCBI Description

2651 Seq. No. Contig ID 1820 1.R1040 dkc700968063.h1 5'-most EST

2652 Seq. No. 1821 1.R1040 Contig ID 5'-most EST LIB3028-052-Q1-B2-E9

2653 Seq. No. 1822 1.R1040 Contig ID

LIB3170-002-Q1-K1-E11 5'-most EST



```
Method
                  BLASTX
NCBI GI
                  q3334663
BLAST score
                  262
E value
                  8.0e-23
Match length
                  114
% identity
                  48
                  (Y10491) putative cytochrome P450 [Glycine max]
NCBI Description
Seq. No.
                  2654
Contig ID
                  1822 2.R1040
5'-most EST
                  LIB3028-052-Q1-B1-F10
Method
                  BLASTX
NCBI GI
                  q3334663
BLAST score
                  186
E value
                   5.0e-14
Match length
                  112
% identity
                  39
NCBI Description
                  (Y10491) putative cytochrome P450 [Glycine max]
Seq. No.
                  2655
                  1823 1.R1040
Contig ID
5'-most EST
                  LIB3170-046-Q1-J1-F3
Method
                  BLASTX
NCBI GI
                  q2239262
BLAST score
                  1028
E value
                  1.0e-111
Match length
                  455
% identity
                   46
                  (Y13285) pectin methylesterase-like protein [Zea mays]
NCBI Description
Seq. No.
                  2656
Contig ID
                  1823 2.R1040
5'-most EST
                  LIB3030-002-Q1-B1-F7
                  2657
Seq. No.
Contig ID
                  1825 1.R1040
5'-most EST
                  LIB3028-052-Q1-B2-F5
Method
                  BLASTX
NCBI GI
                  q418754
BLAST score
                  445
                  5.0e-44
E value
Match length
                  195
% identity
                  55
NCBI Description catechol oxidase (EC 1.10.3.1) precursor - fava bean
                  2658
Seq. No.
                  1827 1.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy118b04b1
Method
                  BLASTX
NCBI GI
                  g833835
BLAST score
                  1372
```

Method BLASTX
NCBI GI 9833835
BLAST score 1372
E value 1.0e-152
Match length 480
% identity 54

NCBI Description (U26025) amygdalin hydrolase isoform AH I precursor [Prunus

serotinal



```
Seq. No.
Contig ID
```

1827 2.R1040

jC-gmle01810014g06d1 5'-most EST

BLASTX Method NCBI GI g2213626 BLAST score 615 1.0e-63 E value 253 Match length % identity

NCBI Description (AC000103) F21J9.18 [Arabidopsis thaliana]

Seq. No.

1827 3.R1040 Contig ID smc700749834.h1 5'-most EST

Seq. No. 2661

Contig ID 1828 1.R1040

5'-most EST LIB3050-018-Q1-E1-D3

Method BLASTX q3334113 NCBI GI 388 BLAST score 3.0e-37 E value 87 Match length % identity

ACYL-COA-BINDING PROTEIN (ACBP) >gi 1006831 (U35015) NCBI Description

acyl-CoA-binding protein [Gossypium hirsutum]

Seq. No.

1829 1.R1040 Contig ID 5'-most EST asn701133385.hl

Method BLASTN NCBI GI g1162979 BLAST score 264 1.0e-146 E value 692 Match length 85 % identity

Spinacia oleracea nuclear-encoded chloroplast NCBI Description

ribulose-5-phosphate 3-epimerase mRNA, complete cds

2663 Seq. No.

1829 2.R1040 Contig ID

jC-gmst02400014f01d1 5'-most EST

BLASTN Method g3264789 NCBI GI 81 BLAST score 2.0e-37 E value Match length 161 88 % identity

Expression vector pFL505 for Spinacia oleracea chloroplast NCBI Description

ribulose-phosphate 3-epimerase transit form

Seq. No. 2664

1830 1.R1040 Contig ID

LIB3170-040-Q1-K1-G3 5'-most EST

Method BLASTX g2407800 NCBI GI 584 BLAST score



```
3.0e-60
E value
Match length
                  134
                  87
% identity
                  (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
NCBI Description
                  2665
Seq. No.
                  1830 2.R1040
Contig ID
                  leu701156037.hl
5'-most EST
                  BLASTX
Method
                  g2407800
NCBI GI
                  179
BLAST score
                   3.0e-13
E value
                   62
Match length
                   61
% identity
NCBI Description (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
                   2666
Seq. No.
                   1831 1.R1040
Contig ID
                  LIB3049-002-Q1-E1-H1
5'-most EST
                  BLASTN
Method
                   g2598656
NCBI GI
                   304
BLAST score
                   1.0e-170
E value
                   551
Match length
                   89
% identity
NCBI Description Vicia faba mRNA for elongation factor 1-alpha (EF1-a)
                   2667
Seq. No.
                   1833 1.R1040
Contig ID
5'-most EST
                   LIB3106-006-Q1-K1-D6
Method
                   BLASTX
                   q2262170
NCBI GI
BLAST score
                   1117
                   1.0e-122
E value
                   350
Match length
% identity
                   (AC002329) predicted glycosyl hydrolase [Arabidopsis
NCBI Description
                   thaliana]
                   2668
Seq. No.
                   1833_2.R1040
Contig ID
                   ssr700558633.hl
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2262170
BLAST score
                   423
E value
                   2.0e-41
Match length
                   183
% identity
                   51
                   (AC002329) predicted glycosyl hydrolase [Arabidopsis
NCBI Description
                   thaliana]
                   2669
Seq. No.
```

Contig ID 5'-most EST 1833 8.R1040 gsv701045546.h1

Seq. No.

2670

Contig ID

1834 1.R1040



uC-gmflminsoy061e09b1 5'-most EST BLASTX Method g1171161 NCBI GI BLAST score 1643 0.0e+00E value 439 Match length 70 % identity NCBI Description (U41472) pectate lyase homolog [Medicago sativa] 2671 Seq. No. 1834 2.R1040 Contig ID zhf700953796.h1 5'-most EST BLASTX Method g1171161 NCBI GI 286 BLAST score 2.0e-25 E value 102 Match length 59 % identity NCBI Description (U41472) pectate lyase homolog [Medicago sativa] Seq. No. 2672 1834 3.R1040 Contig ID LIB3028-052-Q1-B2-B3 5'-most EST BLASTX Method g1171161 NCBI GI 383 BLAST score 9.0e-37 E value 103 Match length 68 % identity (U41472) pectate lyase homolog [Medicago sativa] NCBI Description Seq. No. Contig ID 1835 1.R1040 LIB3170-048-Q1-J1-D7 5'-most EST BLASTX Method q4204260 NCBI GI BLAST score 172 E value 8.0e-12 62 Match length % identity 50 (AC005223) 25568 [Arabidopsis thaliana] NCBI Description Seq. No. 2674 1835 2.R1040 Contig ID uC-gmropic045h01b1 5'-most EST Method BLASTX g4204260 NCBI GI 185 BLAST score 1.0e-13 E value 64 Match length 53 % identity

(AC005223) 25568 [Arabidopsis thaliana] NCBI Description

2675 Seq. No.

1835 4.R1040 Contig ID 5'-most EST fua701041703.hl



```
      Seq. No.
      2676

      Contig ID
      1836_1.R1040

      5'-most EST
      zhf700954832.h1

      Method
      BLASTX
```

NCBI GI g1495251
BLAST score 2468
E value 0.0e+00
Match length 694
% identity 66

NCBI Description (Z70314) heat-shock protein [Arabidopsis thaliana]

Seq. No. 2677

Contig ID 1836_2.R1040

5'-most EST jC-gmro02910048e12d1

Seq. No. 2678

Contig ID 1837_1.R1040

5'-most EST LIB3028-052-Q1-B2-B7

Seq. No. 2679

Contig ID 1839_1.R1040 5'-most EST seb700650924.h1

Method BLASTX
NCBI GI g2950472
BLAST score 185
E value 1.0e-13
Match length 85
% identity 45

NCBI Description (AL022070) putative autophagocytosis protein

[Schizosaccharomyces pombe]

Seq. No. 2680

Contig ID 1842_1.R1040 5'-most EST wvk700682469.h2

Method BLASTN
NCBI GI g1237081
BLAST score 336
E value 0.0e+00
Match length 548
% identity 90

NCBI Description P.sativum mRNA for ADP-glucose pyrophosphorylase (agpS2)

Seq. No. 2681

Contig ID 1842 2.R1040 5'-most EST kmv700743465.h1

Method BLASTX
NCBI GI g1237080
BLAST score 165
E value 1.0e-11
Match length 68
% identity 62

NCBI Description (X96764) ADP-glucose pyrophosphorylase [Pisum sativum]

Seq. No. 2682

Contig ID 1843_1.R1040 5'-most EST asn701138640.h1

Method BLASTX

```
q1885310
NCBI GI
                   298
BLAST score
                   6.0e-27
E value
                   99
Match length
% identity
NCBI Description (X91659) Endoxyloglucan transferase (EXT) [Hordeum vulgare]
                   2683
Seq. No.
                   1846 1.R1040
Contig ID
                   jC-gmst02400073h02d1
5'-most EST
                   2684
Seq. No.
                   1846 2.R1040
Contig ID
                   LIB3028-052-Q1-B2-D10
5'-most EST
                   2685
Seq. No.
                   1847 1.R1040
Contig ID
                   jC-gmle01810012d07a1
5'-most EST
                   BLASTX
Method
                   g2335108
NCBI GI
                   1084
BLAST score
                   1.0e-118
E value
                   307
Match length
                   73
% identity
NCBI Description (AC002339) putative isulinase [Arabidopsis thaliana]
                   2686
Seq. No.
                   1848 1.R1040
Contig ID
                   LIB3107-082-Q1-K1-B11
5'-most EST
                   BLASTX
Method
                   q400515
NCBI GI
                    225
BLAST score
                    3.0e-18
E value
                    90
Match length
                    51
% identity
                    NADH-UBIQUINONE OXIDOREDUCTASE B8 SUBUNIT (COMPLEX I-B8)
NCBI Description
                    (CI-B8) >gi_346540_pir__S28249 NADH dehydrogenase
                    (ubiquinone) (EC 1.6.5.3) chain CI-B8 - bovine
                    >gi_246_emb_CAA44904_ (X63219) NADH dehydrogenase [Bos
                    taurus]
                    2687
Seq. No.
                    1848 2.R1040
Contig ID
                    LIB3039-040-Q1-E1-C7
 5'-most EST
                    BLASTX
Method
                    g400515
NCBI GI
                    226
BLAST score
                    2.0e-18
E value
                    89
Match length
                    49
 % identity
                    NADH-UBIQUINONE OXIDOREDUCTASE B8 SUBUNIT (COMPLEX I-B8)
 NCBI Description
                    (CI-B8) >gi_346540_pir__ S28249 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain CI-B8 - bovine
```

Seq. No.

2688

taurus]

>gi 246_emb_CAA44904_ (X63219) NADH dehydrogenase [Bos



1850 1.R1040 Contig ID $leu7\overline{0}1144978.h1$ 5'-most EST BLASTX Method g4510373 NCBI GI BLAST score 334 E value 5.0e-31 Match length 133 45 % identity (AC007017) putative harpin-induced protein [Arabidopsis NCBI Description thaliana] 2689 Seq. No. 1851 1.R1040 Contig ID leu701155612.h1 5'-most EST 2690 Seq. No. 1855 1.R1040 Contig ID 5'-most EST LIB3049-005-Q1-E1-D1 2691 Seq. No. 1855 2.R1040 Contig ID 5'-most EST kmv700741440.h1 BLASTX Method g2894603 NCBI GI 352 BLAST score 1.0e-32 E value 129 Match length % identity (AL021889) putative protein [Arabidopsis thaliana] NCBI Description Seq. No. 2692 Contig ID 1855 3.R1040 LIB3074-004-Q1-K1-C9 5'-most EST 2693 Seq. No. 1855 6.R1040 Contig ID LIB3092-037-Q1-K1-D2 5'-most EST 2694 Seq. No. 1855 7.R1040 Contig ID LIB3106-010-Q1-K1-C8 5'-most EST Seq. No. 2695 1855 8.R1040 Contig ID uC-qmflminsoy010e06b1 5'-most EST BLASTX Method g2894603 NCBI GI 233 BLAST score

4.0e-19 E value 77 Match length % identity

(AL021889) putative protein [Arabidopsis thaliana] NCBI Description

2696 Seq. No.

1855 11.R1040 Contig ID 5'-most EST smc700744941.h1

BLASTX Method



```
q2894603
NCBI GI
                  189
BLAST score
                  1.0e-14
E value
                  56
Match length
% identity
                  (AL021889) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  1856 1.R1040
Contig ID
                  LIB3040-059-Q1-E1-H6
5'-most EST
Method
                  BLASTX
                  q3786005
NCBI GI
                  845
BLAST score
E value
                   2.0e-90
Match length
                   363
% identity
NCBI Description
                   (AC005499) putative phosphoethanolamine
                  cytidylyltransferase [Arabidopsis thaliana]
Seq. No.
                   1856 2.R1040
Contig ID
5' most EST
                   jex7\overline{0}0909340.h1
                   BLASTX
Method
                   q3786005
NCBI GI
BLAST score
                   584
                   4.0e-60
E value
Match length
                   191
% identity
                   (AC005499) putative phosphoethanolamine
NCBI Description
                   cytidylyltransferase [Arabidopsis thaliana]
                   2699
Seq. No.
Contig ID
                   1857 1.R1040
5'-most EST
                   pcp700991629.hl
Seq. No.
                   2700
Contig ID
                   1859 1.R1040
                   LIB3106-077-P1-K1-G5
5'-most EST
Method
                   BLASTX
                   g1706082
NCBI GI
BLAST score
                   379
                   3.0e-36
E value
Match length
                   131
                   60
% identity
                   SERINE CARBOXYPEPTIDASE II-3 PRECURSOR (CP-MII.3)
NCBI Description
                   >gi 629787 pir S44191 serine-type carboxypeptidase (EC
                   3.4.16.1) II-3 - barley >gi_619350_bbs_153536
                   CP-MII.3=serine carboxypeptidase [Hordeum vulgare=barley,
                   cv. Alexis, aleurone, Peptide, 516 aa]
                   >gi_474392_emb_CAA55478_ (X78877) serine carboxylase II-3
                   [Hordeum vulgare]
Seq. No.
                   2701
                   1860 1.R1040
Contig ID
```

5'-most EST LIB3087-009-Q1-K1-G2

Method BLASTX NCBI GI g4539454



```
276
BLAST score
                   6.0e-24
E value
                   208
Match length
                   37
% identity
NCBI Description (AL049500) contains EST gb:AA728416 [Arabidopsis thaliana]
                   2702
Seq. No.
                   1860 2.R1040
Contig ID
                   LIB3072-052-Q1-E1-F8
5'-most EST
                   2703
Seq. No.
                   1861 1.R1040
Contig ID
                   jex700904613.h1
5'-most EST
                   2704
Seq. No.
                   1862 1.R1040
Contig ID
                   LIB3\overline{1}07-024-Q1-K1-A10
5'-most EST
                   2705
Seq. No.
                   1864 1.R1040
Contig ID
                   LIB3028-052-Q1-B1-B12
5'-most EST
                   2706
Seq. No.
                   1864 2.R1040
Contig ID
                   seb700647923.hl
5'-most EST
                   BLASTX
Method
                   g1439625
NCBI GI
                   174
BLAST score
                   1.0e-12
E value
                   91
Match length
                   37
% identity
                   (U64598) weakly similar to S. cervisiae PTM1 precursor
NCBI Description
                    (SP:P32857) [Caenorhabditis elegans]
                    2707
Seq. No.
                    1865 1.R1040
Contig ID
5'-most EST
                    jC-gmle01810028a08a1
                    2708
Seq. No.
                    1869 1.R1040
Contig ID
                    uC-gmrominsoy283c08b1
 5'-most EST
                    BLASTX
Method
NCBI GI
                    q4432869
BLAST score
                    552
                    1.0e-56
E value
                    157
Match length
 % identity
                   (AC006300) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                    2709
 Seq. No.
                    1869_2.R1040
 Contig ID
```

LIB3106-113-Q1-K1-E5 5'-most EST

Method BLASTX g4432869 NCBI GI 362 BLAST score 4.0e-34 E value 113 Match length



% identity

NCBI Description (AC006300) hypothetical protein [Arabidopsis thaliana]

Seq. No.

Contig ID 5'-most EST 1877 1.R1040 awf700842923.h1

Seq. No.

2711

Contig ID

1877 2.R1040

5'-most EST

LIB3072-060-Q1-K1-F5

Seq. No.

2712

Contig ID

1882 1.R1040

5'-most EST

LIB3055-013-Q1-N1-A7

Seq. No.

2713

Contig ID 5'-most EST 1884 1.R1040 rlr700898336.h1

Method NCBI GI BLASTX g3128177 258

BLAST score E value Match length

2.0e-22 102

% identity

NCBI Description (AC004521) unknown protein [Arabidopsis thaliana]

Seq. No.

2714

46

Contig ID

1885 1.R1040

5'-most EST

LIB3049-027-Q1-E1-F4

Method NCBI GI BLAST score E value

q266945 893 2.0e-96

192

BLASTX

Match length % identity

NCBI Description

60S RIBOSOMAL PROTEIN L9 (GIBBERELLIN-REGULATED PROTEIN GA) >gi 100065_pir__S19978 ribosomal protein L9 - garden pea

>gi 20727 emb CAA46273 (X65155) GA [Pisum sativum] >gi 1279645 emb CAA65987 (X97322) ribosomal protein L9

[Pisum sativum]

Seq. No.

2715

Contig ID 5'-most EST 1885 2.R1040 sat701009469.h1

Method NCBI GI BLASTX g266945 874

BLAST score E value Match length

3.0e - 94192

2716

% identity NCBI Description 88 60S RIBOSOMAL PROTEIN L9 (GIBBERELLIN-REGULATED PROTEIN GA)

>gi_100065_pir__S19978 ribosomal protein L9 - garden pea >gi_20727_emb_CAA46273_ (X65155) GA [Pisum sativum]

>gi 1279645 emb CAA65987 (X97322) ribosomal protein L9

[Pisum sativum]

Seq. No.



```
1885 3.R1040
Contig ID
                  LIB3065-010-Q1-N1-G1
5'-most EST
                  BLASTX
Method
                  g266945
NCBI GI
                  832
BLAST score
                  3.0e-89
E value
Match length
                  192
                  83
% identity
                  60S RIBOSOMAL PROTEIN L9 (GIBBERELLIN-REGULATED PROTEIN GA)
NCBI Description
                  >qi 100065 pir S19978 ribosomal protein L9 - garden pea
                  >gi 20727_emb_CAA46273_ (X65155) GA [Pisum sativum]
                  >gi 1279645_emb_CAA65987_ (X97322) ribosomal protein L9
                   [Pisum sativum]
                  2717
Seq. No.
                  1885 4.R1040
Contig ID
5'-most EST
                  leu701145234.h1
                  BLASTX
Method
                   g2058273
NCBI GI
                   435
BLAST score
                   6.0e-43
E value
                   96
Match length
                   89
% identity
                   (D83527) YK426 [Oryza sativa]
NCBI Description
                   2718
Seq. No.
                   1886 1.R1040
Contig ID
                   uaw700661994.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3790587
BLAST score
                   691
                   1.0e-72
E value
Match length
                   254
                   56
% identity
                   (AF079182) RING-H2 finger protein RHF2a [Arabidopsis
NCBI Description
                   thaliana]
                   2719
Seq. No.
                   1886 2.R1040
Contig ID
                   crh700850064.h1
5'-most EST
Method
                   BLASTX
                   g3790587
NCBI GI
BLAST score
                   313
                   3.0e-28
E value
                   101
Match length
                   64
% identity
                   (AF079182) RING-H2 finger protein RHF2a [Arabidopsis
NCBI Description
                   thaliana]
                   2720
Seq. No.
                   1886 3.R1040
Contig ID
                   LIB3109-005-Q1-K1-H4
```

5'-most EST

Method

BLASTX

NCBI GI BLAST score g3790587 300

E value

3.0e-27

Match length

66



% identity 77
NCBI Description (AF079182) RING-H2 finger protein RHF2a [Arabidopsis thaliana]

Seq. No. 2721

Contig ID 1886_5.R1040

5'-most EST LIB3107-079-Q1-K1-C9

Method BLASTX
NCBI GI g3790587
BLAST score 590
E value 6.0e-61
Match length 167
% identity 66

NCBI Description (AF079182) RING-H2 finger protein RHF2a [Arabidopsis

thaliana]

Seq. No. 2722

Contig ID 1886_6.R1040 5'-most EST sat701010702.h1

Method BLASTX
NCBI GI g3790587
BLAST score 361
E value 2.0e-34
Match length 72
% identity 82

NCBI Description (AF079182) RING-H2 finger protein RHF2a [Arabidopsis

thaliana]

Seq. No. 2723

Contig ID 1888_1.R1040

5'-most EST jC-gmro02910059b02d1

Method BLASTX
NCBI GI 94455158
BLAST score 1148
E value 1.0e-126
Match length 255
% identity 83

NCBI Description (AL021687) kinase-like protein [Arabidopsis thaliana]

Seq. No. 2724

Contig ID 1888_2.R1040
5'-most EST g5175304
Method BLASTX
NCBI GI g3885884
BLAST score 795
E value 5.0e-85
Match length 164

NCBI Description (AF093630) 60S ribosomal protein L21 [Oryza sativa]

Seq. No.

% identity

2725 1888 3 R104

Contig ID 1888 3.R1040

5'-most EST LIB3051-115-Q1-K1-E12

90

Seq. No. 2726

Contig ID 1888_4.R1040 5'-most EST eep700869342.h1



BLASTX Method q4454449 NCBI GI 176 BLAST score 3.0e-12 E value 88 Match length % identity NCBI Description (AC006234) hypothetical protein [Arabidopsis thaliana] 2727 Seq. No. 1888 5.R1040 Contig ID 5'-most EST uC-qmflminsoy034b06b1 2728 Seq. No. 1888 6.R1040 Contig ID g4302443 5'-most EST BLASTX Method NCBI GI g3885884 789 BLAST score 3.0e-84 E value 164 Match length 89 % identity (AF093630) 60S ribosomal protein L21 [Oryza sativa] NCBI Description 2729 Seq. No. 1888 7.R1040 Contig ID rca700998155.h1 5'-most EST BLASTX Method g4455158 NCBI GI 810 BLAST score 2.0e-86 E value 208 Match length % identity (AL021687) kinase-like protein [Arabidopsis thaliana] NCBI Description 2730 Seq. No. 1888 8.R1040 Contig ID 5'-most EST uC-gmrominsoy217c02b1 Seq. No. 2731 1888 9.R1040 Contig ID 5'-most EST jC-gmle01810073h09a1 Seq. No. 2732 1888 10.R1040 Contiq ID pxt700942669.h1 5'-most EST 2733 Seq. No. 1888 11.R1040 Contig ID pmv700892913.h1 5'-most EST 2734 Seq. No. 1888 13.R1040 Contig ID LIB3139-035-P1-N1-G4 5'-most EST

BLASTX Method

g3885884 NCBI GI 501 BLAST score E value 7.0e-51



Match length 104 % identity 88

NCBI Description (AF093630) 60S ribosomal protein L21 [Oryza sativa]

Seq. No. 2735

Contig ID 1891_1.R1040 5'-most EST leu701146963.h1

Method BLASTX
NCBI GI g2088651
BLAST score 250
E value 8.0e-32
Match length 222
% identity 42

NCBI Description (AF002109) hypersensitivity-related gene 201 isolog

[Arabidopsis thaliana]

Seq. No. 2736

Contig ID 1891_2.R1040

5'-most EST LIB3170-006-Q1-K1-H3

Method BLASTN
NCBI GI g2264316
BLAST score 50
E value 8.0e-19
Match length 137
% identity 88

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MRO11, complete sequence [Arabidopsis thaliana]

 Seq. No.
 2737

 Contig ID
 1891_3.R1040

 5'-most EST
 g5753333

 Method
 BLASTX

 NCBI GI
 g2088651

 BLAST score
 169

 E value
 1.0e-11

 Match length
 57

Match length 57 % identity 51

NCBI Description (AF002109) hypersensitivity-related gene 201 isolog

[Arabidopsis thaliana]

Seq. No. 2738

Contig ID 1892 1.R1040

5'-most EST uC-gmronoir032g02b1

Method BLASTX
NCBI GI g4006899
BLAST score 410
E value 4.0e-40
Match length 112
% identity 63

NCBI Description (Z99708) putative protein [Arabidopsis thaliana]

Seq. No. 2739

Contig ID 1892_2.R1040
5'-most EST g4294695
Method BLASTX
NCBI GI g4006899
BLAST score 203



```
E value
                   6.0e-16
Match length
                   64
% identity
                   (Z99708) putative protein [Arabidopsis thaliana]
NCBI Description
                  2740
Seq. No.
                  1892 4.R1040
Contig ID
5'-most EST
                  leu701155723.h1
                  BLASTX
Method
                   q4006899
NCBI GI
BLAST score
                   619
                   1.0e-64
E value
                  156
Match length
                   71
% identity
NCBI Description (Z99708) putative protein [Arabidopsis thaliana]
                   2741
Seq. No.
                   1892 5.R1040
Contig ID
                   fua701041271.hl
5'-most EST
Method
                   BLASTX
                   g4006899
NCBI GI
BLAST score
                   166
                   1.0e-11
E value
                   56
Match length
% identity
                  (Z99708) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1893 1.R1040
Contig ID
5'-most EST
                   uC-qmflminsoy031g02b1
Method
                   BLASTX
                   q2997591
NCBI GI
BLAST score
                   201
                   2.0e-15
E value
                   63
Match length
                   71
% identity
                   (AF020814) glucose-6-phosphate/phosphate-translocator
NCBI Description
                   precursor [Pisum sativum]
                   2743
Seq. No.
Contig ID
                   1893 2.R1040
5'-most EST
                   bth700848169.h1
Method
                   BLASTX
                   q2997591
NCBI GI
                   212
BLAST score
                   8.0e-17
E value
                   68
Match length
% identity
                   71
                   (AF020814) glucose-6-phosphate/phosphate-translocator
NCBI Description
                   precursor [Pisum sativum]
                   2744
Seq. No.
                   1894_1.R1040
Contig ID
                   tku700646334.h1
5'-most EST
```

BLASTX Method g124713 NCBI GI BLAST score 3003



E value 0.0e+00 Match length 645 % identity 87

NCBI Description ACID BETA-FRUCTOFURANOSIDASE PRECURSOR (ACID

SUCROSE-6-PHOSPHATE HYDROLASE) (ACID INVERTASE) (AI) (VACUOLAR INVERTASE) >gi_218326_dbj_BAA01107_ (D10265) invertase [Vigna radiata] >gi_384325_prf__1905412A acid

invertase [Vigna radiata]

Seq. No. 2745

Contig ID 1894_2.R1040 5'-most EST sat701011421.h1

Seq. No. 2746

Contig ID 1894_3.R1040

5'-most EST LIB3051-059-Q1-K2-E11

Method BLASTN
NCBI GI g3821780
BLAST score 36
E value 2.0e-10
Match length 36
% identity 100

NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No. 2747

Contig ID 1894_4.R1040

5'-most EST jC-gmf102220132ad08a1

Method BLASTX
NCBI GI g124713
BLAST score 308
E value 4.0e-28
Match length 115
% identity 54

NCBI Description ACID BETA-FRUCTOFURANOSIDASE PRECURSOR (ACID

SUCROSE-6-PHOSPHATE HYDROLASE) (ACID INVERTASE) (AI) (VACUOLAR INVERTASE) >gi_218326_dbj_BAA01107_ (D10265) invertase [Vigna radiata] >gi_384325_prf__1905412A acid

invertase [Vigna radiata]

Seq. No. 2748

Contig ID 1894 5.R1040

5'-most EST jC-gmf102220138b01d1

Method BLASTN
NCBI GI g218325
BLAST score 229
E value 1.0e-126
Match length 285
% identity 95

NCBI Description Vigna radiata mRNA for invertase, complete cds

Seq. No. 2749

Contig ID 1894 6.R1040

5'-most EST LIB3170-081-Q1-K1-A1

Seq. No. 2750

Contig ID 1896_1.R1040

5'-most EST LIB3109-042-Q1-K1-H10



```
2751
Seq. No.
                   1898 1.R1040
Contig ID
                   leu7\overline{0}1149544.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2129636
                   328
BLAST score
                   4.0e-30
E value
                   219
Match length
                   39
% identity
NCBI Description lipase - Arabidopsis thaliana >gi_1145627 (U38916) lipase
                   [Arabidopsis thaliana]
                   2752
Seq. No.
                   1899 1.R1040
Contig ID
                   bth700844448.h1
5'-most EST
Method
                   BLASTX
                   q4263048
NCBI GI
                   622
BLAST score
                   1.0e-64
E value
                   209
Match length
                   61
% identity
NCBI Description (AC005142) putative hydrolase [Arabidopsis thaliana]
                   2753
Seq. No.
                   1899 2.R1040
Contig ID
5'-most EST
                   fua701038150.h1
                   BLASTX
Method
                   q4263048
NCBI GI
BLAST score
                   278
                   1.0e-24
E value
                   108
Match length
% identity
                   (AC005142) putative hydrolase [Arabidopsis thaliana]
NCBI Description
                   2754
Seq. No.
Contig ID
                   1900 1.R1040
                   LIB3107-061-Q1-K1-G1
5'-most EST
Method
                   BLASTN
                   g456567
NCBI GI
BLAST score
                   275
                   1.0e-153
E value
Match length
                   451
                   90
% identity
                   Pisum sativum ubiquitin conjugating enzyme (UBC4), complete
NCBI Description
                   2755
Seq. No.
                   1900 3.R1040
Contig ID
                   LIB3049-023-Q1-E1-B9
5'-most EST
                   BLASTN
Method
                    g456567
NCBI GI
BLAST score
                    67
                    3.0e-29
E value
                    187
Match length
```

NCBI Description Pisum sativum ubiquitin conjugating enzyme (UBC4), complete

84

% identity



2756 Seq. No. 1900 4.R1040 Contig ID bth700846257.h1 5'-most EST

BLASTN Method q456567 NCBI GI BLAST score 65 4.0e-28 E value 178 Match length

90 % identity Pisum sativum ubiquitin conjugating enzyme (UBC4), complete NCBI Description

cds

2757 Seq. No.

1901 1.R1040 Contig ID ujr700646501.h1 5'-most EST

BLASTN Method g4193381 NCBI GI 85 BLAST score 8.0e-40 E value Match length 261 83 % identity

Arabidopsis thaliana ribosomal protein S27 (ARS27A) mRNA, NCBI Description

complete cds

2758 Seq. No.

1901 2.R1040 Contiq ID

LIB3106-088-Q1-K1-G8 5'-most EST

BLASTN Method g4193381 NCBI GI BLAST score 91 2.0e-43 E value 239 Match length

85 % identity

Arabidopsis thaliana ribosomal protein S27 (ARS27A) mRNA, NCBI Description

complete cds

2759 Seq. No.

1901 3.R1040 Contig ID q5606680 5'-most EST Method BLASTX NCBI GI q4193382 BLAST score 367 4.0e-35 E value 83 Match length

% identity (AF083336) ribosomal protein S27 [Arabidopsis thaliana] NCBI Description

>gi 4193384 (AF083337) ribosomal protein S27 [Arabidopsis

thaliana]

2760 Seq. No.

1902 1.R1040 Contig ID

LIB3049-016-Q1-E1-D12 5'-most EST

2761 Seq. No.

1902 2.R1040 Contig ID



5'-most EST bth700849686.h1

Seq. No. 2762

Contig ID 1902 3.R1040

5'-most EST LIB3107-006-Q1-K1-C9

Seq. No. 2763

Contig ID 1902 5.R1040

5'-most EST LIB3139-006-P1-N1-H2

Seq. No. 2764 Contig ID 1905 1.R1040

5'-most EST LIB3049-011-Q1-E1-F10

Method BLASTX
NCBI GI g3608481
BLAST score 610
E value 3.0e-63
Match length 146
% identity 78

NCBI Description (AF088913) ribosomal protein L27a [Petunia x hybrida]

Seq. No. 2765

Contig ID 1905_3.R1040 5'-most EST txt700735894.h1

Method BLASTN
NCBI GI g1881580
BLAST score 60
E value 4.0e-25
Match length 124
% identity 88

NCBI Description Drosophila melanogaster ribosomal protein RpL27a gene,

complete cds

Seq. No. 2766

Contig ID 1906_1.R1040

5'-most EST LIB3028-051-Q1-B1-F9

Seq. No. 2767

Contig ID 1907_1.R1040 5'-most EST kl1701208895.h1

Method BLASTX
NCBI GI g2583108
BLAST score 643
E value 9.0e-67
Match length 311
% identity 45

NCBI Description (AC002387) putative surface protein [Arabidopsis thaliana]

Seq. No.

2768

Contig ID 1907 2.R1040

5'-most EST jC-gmst02400070b04d1

Seq. No. 2769

Contig ID 1908_1.R1040 5'-most EST fua701042926.h1

Method BLASTX NCBI GI g2129648



BLAST score 729 E value 6.0e-77 Match length 295 % identity 55

NCBI Description MYB-related protein 33,3K - Arabidopsis thaliana

>gi_1263095_emb_CAA90809_ (Z54136) MYB-related protein

[Arabidopsis thaliana]

Seq. No. Contig ID 2770 1908 2.R1040

5'-most EST

LIB3028-013-Q1-B1-C2

Method BLASTX
NCBI GI g2129648
BLAST score 223
E value 8.0e-18
Match length 177
% identity 36

NCBI Description MYB-related protein 33,3K - Arabidopsis thaliana

>gi 1263095 emb CAA90809 (Z54136) MYB-related protein

[Arabidopsis thaliana]

Seq. No.

Contig ID 5'-most EST 1908_3.R1040 crh700852179.h1

Method BLASTX
NCBI GI g2642435
BLAST score 559
E value 1.0e-57
Match length 112
% identity 89

NCBI Description (AC002391) MYB-related protein [Arabidopsis thaliana]

Seq. No.

2772

2771

Contig ID 5'-most EST 1909_1.R1040 hrw701061638.h1

Method BLASTN
NCBI GI g1326162
BLAST score 340
E value 0.0e+00
Match length 536
% identity 91

NCBI Description Phaseolus vulgaris stress related protein PvSRP mRNA,

complete cds

Seq. No.

2773

2774

Contig ID

1909 2.R1040

5'-most EST

LIB3051-043-Q1-K1-A4

Method BLASTN
NCBI GI g1326162
BLAST score 324
E value 0.0e+00
Match length 536
% identity 90

NCBI Description Phaseolus vulgaris stress related protein PvSRP mRNA,

complete cds

Seq. No.



```
1909 3.R1040
Contig ID
                  LIB3093-040-Q1-K1-F9
5'-most EST
                  BLASTN
Method
                  g1326162
NCBI GI
                  139
BLAST score
                  6.0e-72
E value
Match length
                  267
                  88
% identity
                  Phaseolus vulgaris stress related protein PvSRP mRNA,
NCBI Description
                  complete cds
                  2775
Seq. No.
                  1912 1.R1040
Contig ID
                  LIB3107-075-Q1-K1-C9
5'-most EST
                  BLASTX
Method
                   g3077640
NCBI GI
                   852
BLAST score
                   1.0e-91
E value
                   268
Match length
                   63
% identity
NCBI Description (AJ223151) O-methyltransferase [Prunus dulcis]
                   2776
Seq. No.
                   1912 2.R1040
Contig ID
                   zhf7\overline{0}0954691.h1
5'-most EST
                   BLASTX
Method
                   g1174621
NCBI GI
BLAST score
                   879
                   1.0e-94
E value
                   330
Match length
                   52
% identity
                   T-COMPLEX PROTEIN 1, THETA SUBUNIT (TCP-1-THETA)
NCBI Description
                   (CCT-THETA) >gi 1083259 pir _JC4073 TCP-1 containing
                   cytosolic chaperonin (CCT) Theta chain - mouse
                   >gi_695625_emb_CAA85521_ (Z37164) CCTtheta, theta subunit
                   of the chaperonin containing TCP-1 (CCT) [Mus musculus]
                   2777
Seq. No.
                   1915 1.R1040
Contig ID
                   zhf700952177.h1
5'-most EST
                   BLASTX
Method
                   g2739389
NCBI GI
BLAST score
                   1268
                   1.0e-140
E value
                   450
Match length
                   57
% identity
                   (AC002505) Cf-2.2 like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2778
                   1916 1.R1040
Contig ID
                    leu7\overline{0}1150759.h1
 5'-most EST
                   BLASTX
Method
NCBI GI
                    g4559334
                    1197
BLAST score
                    1.0e-132
 E value
```

326

63

Match length

% identity

```
NCBI Description (AC007087) unknown protein [Arabidopsis thaliana]
                  2779
Seq. No.
                  1917 1.R1040
Contig ID
                  LIB3109-046-Q1-K1-H11
5'-most EST
                  BLASTX
Method
                  g559330
NCBI GI
BLAST score
                  276
                  8.0e-24
E value
                  306
Match length
                   26
% identity
NCBI Description (D38521) The ha0919 gene product is novel. [Homo sapiens]
                   2780
Seq. No.
                   1918 1.R1040
Contig ID
                   LIB3028-035-Q1-B1-A7
5'-most EST
                   BLASTX
Method
                   g2213629
NCBI GI
                   675
BLAST score
                   7.0e-71
E value
                   190
Match length
                   66
% identity
NCBI Description (AC000103) F21J9.21 [Arabidopsis thaliana]
                   2781
Seq. No.
                   1919 1.R1040
Contia ID
                   LIB3106-049-Q1-K1-H7
5'-most EST
                   BLASTX
Method
                   g4049349
NCBI GI
                   399
BLAST score
                   1.0e-38
E value
                   122
Match length
                   63
% identity
```

(AL034567) ubiquinol-cytochrome c reductase-like protein NCBI Description [Arabidopsis thaliana]

2782 Seq. No. Contig ID 1920 1.R1040

LIB3028-051-Q1-B1-E1 5'-most EST

2783 Seq. No.

Contig ID 1922 1.R1040

5'-most EST LIB3109-022-Q1-K1-A9

Method BLASTX NCBI GI q4432814 305 BLAST score E value 3.0e-27 238 Match length % identity

(AC006593) unknown protein [Arabidopsis thaliana] NCBI Description

2784 Seq. No.

Contig ID 1924 1.R1040 jex700907752.hl 5'-most EST

BLASTX Method q4455192 NCBI GI 240 BLAST score



E value 3.0e-20 Match length 126 % identity 22

NCBI Description (AL035440) putative protein [Arabidopsis thaliana]

Seq. No. 278

Contig ID 1925_1.R1040

5'-most EST uC-gmflminsoy026b09b1

Method BLASTX
NCBI GI 9400890
BLAST score 662
E value 4.0e-69
Match length 247
% identity 60

NCBI Description PHOTOSYSTEM II 22 KD PROTEIN PRECURSOR

>gi_282837_pir__S26953 photosystem II 22K protein precursor
- spinach >gi_21307_emb_CAA48557_ (X68552) 22kD-protein of
PSII [Spinacia oleracea] >gi_260917_bbs_119338 (S49864)
photosystem II 22 kda polypeptide [spinach, Peptide, 274

aa] [Spinacia oleracea]

Seq. No. 2786

Contig ID 1925_2.R1040

5'-most EST jC-gmle01810064a12a1

Method BLASTX
NCBI GI g1709846
BLAST score 285
E value 7.0e-34
Match length 132
% identity 43

NCBI Description PHOTOSYSTEM II 22 KD PROTEIN PRECURSOR >gi_706853 (U04336)

22 kDa component of photosystem II [Lycopersicon

esculentum]

Seq. No. 2787

Contig ID 1925_3.R1040

5'-most EST LIB3138-079-P1-N1-H6

Method BLASTN
NCBI GI g21306
BLAST score 45
E value 5.0e-16
Match length 117
% identity 85

NCBI Description S.oleracea mRNA for photosystem II 22kDa protein

Seq. No. 2788

Contig ID 1925_7.R1040 5'-most EST ncj700981872.h1

Seq. No. 2789

Contig ID 1925_8.R1040 5'-most EST kmv700741481.h1

Seq. No. 2790

Contig ID 1927_1.R1040 5'-most EST kmv700739216.h1

Method BLASTX



NCBI GI g2499553 BLAST score 1797 E value 0.0e+00 Match length 476 % identity 75

NCBI Description CRYPTOCHROME 1 APOPROTEIN (BLUE LIGHT PHOTORECEPTOR)

>gi_629524_pir__S39058 probable deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - Arabidopsis thaliana

>gi_442529_bbs_139743 (S66907) flavin-type blue-light photoreceptor, HY4=DNA photolyase/tropomyosin A homolog [Arabidopsis thaliana, ecotype Columbia, Peptide, 681 aa]

[Arabidopsis thaliana]

Seq. No. 2791

Contig ID 1927_2.R1040 5'-most EST kl1701204912.h1

Method BLASTX
NCBI GI 94325368
BLAST score 271
E value 1.0e-23
Match length 157
% identity 46

NCBI Description (AF128396) Arabidopsis thaliana flavin-type blue-light

photoreceptor (SW:Q43125) (Pfam: PF00875, Score=765.2,

E=2.6e-226, N=1) [Arabidopsis thaliana]

Seq. No. 2792

Contig ID 1927 3.R1040

5'-most EST LIB3093-018-Q1-K1-E1

Seq. No. 2793

Contig ID 1929_1.R1040 5'-most EST leu701146395.h1

Method BLASTX
NCBI GI g2222777
BLAST score 301
E value 3.0e-27
Match length 106
% identity 58

NCBI Description (Y13986) GTP-binding protein [Ostertagia circumcincta]

Seq. No. 2794

Contig ID 1930_1.R1040

5'-most EST LIB3028-051-Q1-B1-B4

Method BLASTX
NCBI GI g3426036
BLAST score 239
E value 5.0e-20
Match length 128
% identity 47

NCBI Description (AC005168) unknown protein [Arabidopsis thaliana]

Seq. No.

Contig ID 1931_1.R1040

5'-most EST LIB3028-019-Q1-B1-A3

2795

Seq. No. 2796

Method

NCBI GI

E value

BLAST score

Match length

BLASTX g3860263

1.0e-173

1554

460



```
1933 1.R1040
Contig ID
                  LIB3028-051-Q1-B1-B7
5'-most EST
                  2797
Seq. No.
                  1934 1.R1040
Contig ID
                  LIB3028-051-Q1-B1-B8
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4309969
                  241
BLAST score
                   2.0e-20
E value
                  89
Match length
                   52
% identity
                   (AC002983) putative phosphoglyceride transfer protein
NCBI Description
                   [Arabidopsis thaliana]
                   2798
Seq. No.
                   1934 2.R1040
Contig ID
                   LIB3049-045-Q1-E1-F12
5'-most EST
                   2799
Seq. No.
                   1936 1.R1040
Contig ID
                   leu701155657.h1
5'-most EST
                   BLASTX
Method
                   q100347
NCBI GI
                   383
BLAST score
                   9.0e-37
E value
                   141
Match length
                   53
% identity
                   monosaccharide transport protein MST1 - common tobacco
NCBI Description
                   >gi_19885_emb_CAA47324_ (X66856) monosaccharid transporter
                   [Nicotiana tabacum]
                   2800
Seq. No.
                   1936 2.R1040
Contig ID
                   uC-gmflminsoy079f01b1
5'-most EST
                   BLASTX
Method
                   g100347
NCBI GI
BLAST score
                   312
                   1.0e-28
E value
                   98
Match length
                   59
% identity
                   monosaccharide transport protein MST1 - common tobacco
NCBI Description
                   >gi 19885_emb_CAA47324_ (X66856) monosaccharid transporter
                   [Nicotiana tabacum]
                   2801
Seq. No.
                   1940 1.R1040
Contig ID
                   LIB3170-078-Q1-J1-D1
 5'-most EST
                   2802
 Seq. No.
                   1941 1.R1040
 Contig ID
                   uC-gmrominsoy180d02b1
 5'-most EST
```



% identity (AC005824) putative cytochrome p450 protein [Arabidopsis NCBI Description thaliana] 2803 Seq. No. 1941 2.R1040 Contig ID uC-gmrominsoy243e12b1 5'-most EST Method BLASTX q3860263 NCBI GI 658 BLAST score 5.0e-70 E value 241 Match length 58 % identity (AC005824) putative cytochrome p450 protein [Arabidopsis NCBI Description thaliana] 2804 Seq. No. 1941 4.R1040 Contig ID LIB3092-058-Q1-K1-E8 5'-most EST BLASTX Method g3860263 NCBI GI 288 BLAST score 8.0e-26 E value 91 Match length % identity (AC005824) putative cytochrome p450 protein [Arabidopsis NCBI Description thaliana] 2805 Seq. No. 1942 1.R1040 Contig ID LIB3028-043-Q1-B1-E4 5'-most EST BLASTX Method g2244956 NCBI GI 307 BLAST score 1.0e-27 E value 146 Match length 43 % identity (Z97340) strong similarity to pectinesterase [Arabidopsis NCBI Description thaliana] 2806 Seq. No. 1942 2.R1040 Contig ID 5'-most EST jC-qmf102220130e04d1 2807 Seq. No. 1944 1.R1040 Contig ID 5'-most EST LIB3107-044-Q1-K1-E9 2808 Seq. No. 1944 3.R1040 Contig ID 5'-most EST zhf700964515.hl

Seq. No. 2809

Contig ID 1946_1.R1040 5'-most EST sat701015379.h1

Method BLASTX NCBI GI g2760320



BLAST score 504 E value 2.0e-50 Match length 211 % identity 49

NCBI Description (AC002130) F1N21.4 [Arabidopsis thaliana]

Seq. No. 2810

Contig ID 1946_2.R1040

5'-most EST jC-gmst02400054f03d1

Seq. No. 2811

Contig ID 1946_3.R1040 5'-most EST rlr700897008.h1

Method BLASTX
NCBI GI g2760320
BLAST score 203
E value 3.0e-17
Match length 112
% identity 54

NCBI Description (AC002130) F1N21.4 [Arabidopsis thaliana]

Seq. No. 2812

Contig ID 1946_5.R1040 5'-most EST zsg701118627.h1

Seq. No. 2813

Contig ID 1946_6.R1040 5'-most EST smc700746556.h1

Seq. No. 2814

Contig ID 1947_1.R1040 5'-most EST gsv701046102.h1

Method BLASTN
NCBI GI g463251
BLAST score 245
E value 1.0e-135
Match length 553
% identity 86

NCBI Description M.sativa (Nagyszenasi) mRNA for ribosomal protein RL5

Seq. No. 2815

Contig ID 1947_2.R1040

5'-most EST LIB3040-025-Q1-E1-H8

Method BLASTN
NCBI GI g1944341
BLAST score 340
E value 0.0e+00
Match length 507
% identity 92

NCBI Description Glycine max DNA for cysteine proteinase inhibitor, complete

cds

Seq. No. 2816

Contig ID 1947 3.R1040

5'-most EST uC-gmropic108c03b1

Method BLASTN NCBI GI g463251



162 BLAST score 1.0e-85 E value Match length 354 86 % identity

NCBI Description M.sativa (Nagyszenasi) mRNA for ribosomal protein RL5

Seq. No.

2817

Contig ID 5'-most EST

1947 4.R1040 LIB3028-009-Q1-B1-F12

Method NCBI GI BLASTN g4512656

BLAST score

67

E value Match length 4.0e-29 115

% identity

90

NCBI Description Arabidopsis thaliana chromosome II BAC F7D19 genomic

sequence, complete sequence

Seq. No.

2818

Contig ID

1947 5.R1040

5'-most EST

LIB3170-016-Q1-K1-G1

Method NCBI GI BLAST score BLASTN q463251 241

E value

1.0e-133

Match length % identity

553 86

NCBI Description

M.sativa (Nagyszenasi) mRNA for ribosomal protein RL5

Seq. No.

2819

Contig ID

1947 7.R1040

5'-most EST

LIB3040-019-01-E1-F10

Method NCBI GI BLASTX g1173055

BLAST score

462

E value Match length 3.0e-46124 81

% identity

NCBI Description

60S RIBOSOMAL PROTEIN L11 (L5) >gi 541961 pir S42497 ribosomal protein L11.e - alfalfa >gi 1076504 pir S51819 RL5 ribosomal protein - alfalfa >gi_463252 emb_CAA55090_

(X78284) RL5 ribosomal protein [Medicago sativa]

Seq. No.

2820

Contig ID 5'-most EST 1947 9.R1040 zzp700829957.h1

Method NCBI GI BLASTN g1944341

BLAST score

87

E value Match length

3.0e-41 147

% identity NCBI Description

90 Glycine max DNA for cysteine proteinase inhibitor, complete

Seq. No.

2821

Contig ID

1947 12.R1040

```
5'-most EST
                   gsv701046130.hl
Method
                   BLASTN
                   q463251
NCBI GI
BLAST score
                   102
                   3.0e-50
E value
Match length
                   250
                   85
% identity
NCBI Description M.sativa (Nagyszenasi) mRNA for ribosomal protein RL5
Seq. No.
                   2822
Contig ID
                   1948 1.R1040
5'-most EST
                  LIB3049-010-Q1-E1-A1
Method
                  BLASTX
NCBI GI
                  g3402711
BLAST score
                  779
E value
                   1.0e-82
Match length
                  360
% identity
                   50
NCBI Description
                   (AC004261) putative RNA-binding protein [Arabidopsis
                  thaliana]
Seq. No.
                  2823
Contig ID
                  1948 2.R1040
5'-most EST
                  uC-gmflminsoy045h11b1
Seq. No.
                   2824
                   1948 5.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy043e03b1
Seq. No.
                  2825
Contig ID
                  1949 1.R1040
5'-most EST
                  LIB3065-007-Q1-N1-B3
Method
                  BLASTX
NCBI GI
                  g2244744
BLAST score
                   311
E value
                   5.0e-28
Match length
                  139
% identity
                   50
NCBI Description
                  (Y13676) bZIP DNA-binding protein [Antirrhinum majus]
                  2826
Seq. No.
Contig ID
                  1949 2.R1040
5'-most EST
                  leu701155981.h1
Method
                  BLASTN
NCBI GI
                  g394735
BLAST score
                  44
E value
                  2.0e-15
Match length
                  52
% identity
                  96
NCBI Description
                  Rice lip19 mRNA for basic/leucine zipper protein
```

Seq. No. 2827

Contig ID 1950 1.R1040 5'-most EST ncj700982269.h1 Method BLASTX

Method BLASTX NCBI GI g1778376 BLAST score 1471



0.0e + 00E value Match length 374 89 % identity

(U81288) PsRT17-1 [Pisum sativum] NCBI Description

Seq. No.

2828

Contig ID

1950 2.R1040

5'-most EST

jC-gmst02400050d06a1

Method BLASTN g1778375 NCBI GI BLAST score 50 7.0e-19 E value

86 Match length 90 % identity

NCBI Description Pisum sativum PsRT17-1 mRNA, complete cds

Seq. No.

2829

Contig ID

1950 5.R1040

5'-most EST

LIB3065-006-Q1-N1-E7

Method BLASTN NCBI GI g1778375 BLAST score 36 7.0e-11 E value Match length 74 % identity 93

NCBI Description Pisum sativum PsRT17-1 mRNA, complete cds

Seq. No.

2830

Contig ID 5'-most EST 1950 6.R1040 zhf700962217.h1

Method BLASTN q1778375 NCBI GI BLAST score 159 4.0e-84 E value Match length 377

87 % identity

NCBI Description Pisum sativum PsRT17-1 mRNA, complete cds

Seq. No.

2831

Contig ID 5'-most EST 1950 9.R1040 wrg700786756.h2

Seq. No.

2832

Contig ID

1951 1.R1040

5'-most EST

LIB3139-087-P1-N1-D6

Method BLASTX NCBI GI q2654358 BLAST score 494 E value 1.0e-49 Match length 175 54

% identity NCBI Description

(Y15522) MNUDC protein [Mus musculus]

>gi 2808636_emb_CAA57201_ (X81443) Sig 92 [Mus musculus]

Seq. No.

2833

Contig ID

1951 2.R1040 $leu7\overline{0}1151540.h1$

5'-most EST



```
2834
Seq. No.
                  1952 1.R1040
Contig ID
                   jC-gmro02910062e09a1
5'-most EST
                  BLASTN
Method
                  g18729
NCBI GI
BLAST score
                   36
                   1.0e-10
E value
                   40
Match length
                   97
% identity
NCBI Description Soybean (Glycine max) 18S ribosomal RNA
                   2835
Seq. No.
                   1952 2.R1040
Contig ID
                   LIB3028-051-Q1-B1-A12
5'-most EST
                   BLASTX
Method
                   g3927831
NCBI GI
                   740
BLAST score
                   3.0e-78
E value
                   199
Match length
% identity
                   (AC005727) similar to mouse ankyrin 3 [Arabidopsis
NCBI Description
                   thaliana]
                   2836
Seq. No.
                   1952 3.R1040
Contig ID
                   g5607045
5'-most EST
                   BLASTX
Method
                   g3927831
NCBI GI
                   419
BLAST score
E value
                   6.0e-41
                   163
Match length
% identity
                   (AC005727) similar to mouse ankyrin 3 [Arabidopsis
NCBI Description
                   thaliana]
                   2837
Seq. No.
Contig ID
                   1952 4.R1040
                   r1r7\overline{0}0896407.h1
5'-most EST
                   BLASTX
Method
                   g3927831
NCBI GI
BLAST score
                   336
E value
                    4.0e-31
Match length
                   175
% identity
                    (AC005727) similar to mouse ankyrin 3 [Arabidopsis
NCBI Description
                   thaliana]
                    2838
Seq. No.
                    1955 1.R1040
Contig ID
                   LIB3094-070-Q1-K1-F5
5'-most EST
                   BLASTX
Method
                   g3482979
NCBI GI
                    548
BLAST score
                    3.0e-56
E value
                    127
Match length
```

74

% identity



NCBI Description (AL031369) putative protein [Arabidopsis thaliana] >gi 4567258 gb AAD23672.1 AC007070 21 (AC007070)

hypothetical protein [Arabidopsis thaliana]

Seq. No. 2839

Contig ID 1956_1.R1040 5'-most EST crh700854558.h1

Method BLASTX
NCBI GI g133872
BLAST score 1491
E value 1.0e-166
Match length 363
% identity 80

NCBI Description 30S RIBOSOMAL PROTEIN S1, CHLOROPLAST PRECURSOR (CS1)

>gi_282838_pir__S26494 ribosomal protein S1, chloroplast spinach >gi_322404_pir__A44121 small subunit ribosomal
protein CS1, CS-S2 - spinach >gi_18060_emb_CAA46927_
(X66135) ribosomal protein S1 [Spinacia oleracea]
>gi_170143 (M82923) chloroplast ribosomal protein S1

[Spinacia oleracea]

Seq. No. 2840

Contig ID 1956 2.R1040

5'-most EST LIB3106-006-Q1-K1-G8

Method BLASTX
NCBI GI g133872
BLAST score 469
E value 5.0e-47
Match length 101
% identity 93

NCBI Description 30S RIBOSOMAL PROTEIN S1, CHLOROPLAST PRECURSOR (CS1)

>gi_282838_pir__S26494 ribosomal protein S1, chloroplast spinach >gi_322404_pir__A44121 small subunit ribosomal
protein CS1, CS-S2 - spinach >gi_18060_emb_CAA46927_
(X66135) ribosomal protein S1 [Spinacia oleracea]
>gi_170143 (M82923) chloroplast ribosomal protein S1

[Spinacia oleracea]

Seq. No. 2841

Contig ID 1957_1.R1040 5'-most EST bth700848861.h1

Method BLASTX
NCBI GI g1076610
BLAST score 1404
E value 1.0e-156
Match length 355
% identity 70

NCBI Description cathepsin B-like cysteine proteinase (EC 3.4.22.-) - Aztec

tobacco >qi 609175 emb CAA57522 (X81995) cathepsin B-like

cysteine proteinase [Nicotiana rustica]

Seq. No. 2842

Contig ID 1957 2.R1040

5'-most EST jC-gmf102220056d01d1

Seq. No. 2843

Contig ID 1957 3.R1040

Contig ID

5'-most EST



```
jC-gmle01810009b12a1
5'-most EST
                  BLASTX
Method
                  q1652807
NCBI GI
                  261
BLAST score
                  3.0e-22
E value
Match length
                  114
% identity
NCBI Description (D90908) hypothetical protein [Synechocystis sp.]
Seq. No.
                  1957 4.R1040
Contig ID
                  uC-gmrominsoy176c09b1
5'-most EST
                  BLASTX
Method
                   g2842493
NCBI GI
                   483
BLAST score
                   3.0e-48
E value
                   187
Match length
                   50
% identity
NCBI Description (AL021749) predicted protein [Arabidopsis thaliana]
                   2845
Seq. No.
                   1957 5.R1040
Contig ID
                   iC-qmle01810037h10a2
5'-most EST
                   BLASTN
Method
                   q21952
NCBI GI
                   95
BLAST score
                   8.0e-46
E value
                   175
Match length
% identity
                   T.repens mRNA for cyanogenic beta-glucosidase (linamarase)
NCBI Description
                   2846
Seq. No.
                   1957 6.R1040
Contig ID
                   jC-gmle01810001c03a1
5'-most EST
                   BLASTN
Method
                   g21952
NCBI GI
BLAST score
                   89
                   5.0e-42
E value
                   153
Match length
                   90
 % identity
                   T.repens mRNA for cyanogenic beta-glucosidase (linamarase)
NCBI Description
                   2847
Seq. No.
                   1957 8.R1040
Contig ID
                   uC-gmropic097d08b1
 5'-most EST
                   BLASTX
Method
                   g2842493
 NCBI GI
                   193
 BLAST score
                   1.0e-14
 E value
                   53
 Match length
 % identity
                   (AL021749) predicted protein [Arabidopsis thaliana]
 NCBI Description
                    2848
 Seq. No.
                    1957 12.R1040
```

zsg701129432.h1



```
2849
Seq. No.
                  1958 1.R1040
Contig ID
                  LIB3028-051-Q1-B1-A8
5'-most EST
Method
                  BLASTX
                   g3256035
NCBI GI
                   956
BLAST score
                   1.0e-107
E value
Match length
                   458
% identity
                   (Y14274) putative serine/threonine protein kinase [Sorghum
NCBI Description
                   bicolor]
                   2850
Seq. No.
                   1958 2.R1040
Contig ID
                   fC-gmse700855803a1
5'-most EST
                   BLASTX
Method
                   g2980770
NCBI GI
                   576
BLAST score
E value
                   5.0e-59
                   284
Match length
                   45
% identity
                   (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   1958 4.R1040
Contig ID
                   uC-gmropic107g01b1
5'-most EST
                   2852
Seq. No.
                   1958 5.R1040
Contig ID
                   seb7\overline{0}0653353.h1
5'-most EST
                   2853
Seq. No.
                   1958 6.R1040
Contig ID
                   6HC-01-Q1-B1-E11
5'-most EST
                   BLASTX
Method
                   g1777312
NCBI GI
                   153
BLAST score
                   7.0e-10
E value
                   45
Match length
% identity
                    (D30622) novel serine/threonine protein kinase [Arabidopsis
NCBI Description
                   thaliana]
                    2854
Seq. No.
                    1958 8.R1040
Contig ID
                    jC-qmro02910020g05d1
5'-most EST
                   BLASTX
Method
                   g2980770
NCBI GI
BLAST score
                    183
                    2.0e-13
E value
                    65
Match length
% identity
                    (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
```

595

2855

1959 1.R1040

Seq. No.

Contig ID

5'-most EST

NCBI Description



```
BLASTX
Method
                  q3080391
NCBI GI
                  679
BLAST score
                   5.0e-71
E value
                   292
Match length
                   67
% identity
NCBI Description (AL022603) putative protein [Arabidopsis thaliana]
                   2856
Seq. No.
                   1961 1.R1040
Contig ID
                  wvk700681379.h2
5'-most EST
                  BLASTX
Method
                   g3341899
NCBI GI
                   185
BLAST score
                   4.0e-13
E value
                   110
Match length
                   45
% identity
                  (AB009285) BCNT [Homo sapiens]
NCBI Description
                   2857
Seq. No.
                   1962 1.R1040
Contig ID
                   jC-gmro02800033c10a1
5'-most EST
                   BLASTX
Method
                   q2129956
NCBI GI
                   467
BLAST score
                   2.0e-46
E value
                   155
Match length
% identity
                   photoassimilate-responsive protein PAR-1c precursor -
NCBI Description
                   common tobacco >gi 871489 emb_CAA58732_ (X83852) mRNA
                   inducible by sucrose and salicylic acid and potato virus Y
                   expressed in sugar-accumulating plants [Nicotiana tabacum]
                   2858
Seq. No.
                   1962 2.R1040
Contig ID
                   ncj700982510.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2129956
                   353
BLAST score
                   2.0e-33
E value
                   115
Match length
 % identity
                   photoassimilate-responsive protein PAR-1c precursor -
NCBI Description
                   common tobacco >gi_871489_emb_CAA58732_ (X83852) mRNA
                   inducible by sucrose and salicylic acid and potato virus Y
                   expressed in sugar-accumulating plants [Nicotiana tabacum]
                   2859
 Seq. No.
                   1962 3.R1040
 Contig ID
                   g4437054
 5'-most EST
                   BLASTX
 Method
                   q4544412
 NCBI GI
 BLAST score
                    384
 E value
                    1.0e-36
                   173
 Match length
                    40
 % identity
```

(AC006955) hypothetical protein [Arabidopsis thaliana]



```
2860
Seq. No.
                  1965 1.R1040
Contig ID
                  epx701106858.h1
5'-most EST
                  BLASTX
Method
                  g4388726
NCBI GI
BLAST score
                  1425
                  1.0e-158
E value
                  383
Match length
                  71
% identity
                   (AC006413) putative 12-oxophytodienoate-10,11-reductase
NCBI Description
                   [Arabidopsis thaliana]
                   2861
Seq. No.
                   1965 2.R1040
Contig ID
                  pmv700893580.h1
5'-most EST
                   BLASTX
Method
                   q4388726
NCBI GI
BLAST score
                   811
                   7.0e-87
E value
                   198
Match length
                   76
% identity
                   (AC006413) putative 12-oxophytodienoate-10,11-reductase
NCBI Description
                   [Arabidopsis thaliana]
                   2862
Seq. No.
                   1965 3.R1040
Contig ID
                   LIB3170-037-Q1-J1-C12
5'-most EST
                   2863
Seq. No.
                   1965 4.R1040
Contig ID
                   LIB3109-009-Q1-K1-G10
5'-most EST
                   BLASTX
Method
                   q4388726
NCBI GI
                   390
BLAST score
                   9.0e-38
E value
                   97
Match length
% identity
                   (AC006413) putative 12-oxophytodienoate-10,11-reductase
NCBI Description
                   [Arabidopsis thaliana]
                   2864
Seq. No.
Contig ID
                   1966 1.R1040
                   vzy700756786.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4220512
BLAST score
                   229
                   3.0e-19
E value
Match length
                   57
                   72
% identity
                   (AL035356) putative pectate lyase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2865
```

1967 1.R1040 Contig ID

LIB3028-052-Q1-B2-E1 5'-most EST

BLASTX Method g3608412 NCBI GI



BLAST score 662 E value 2.0e-69 Match length 179 % identity 73

NCBI Description (AF079355) protein phosphatase-2c [Mesembryanthemum

crystallinum]

Seq. No. 2866

Contig ID 1969_1.R1040

5'-most EST LIB3028-040-Q1-B1-A7

Method BLASTX
NCBI GI g4510395
BLAST score 342
E value 3.0e-42
Match length 114
% identity 75

NCBI Description (AC006587) putative beta-galactosidase precursor

[Arabidopsis thaliana]

Seq. No. 2867

Contig ID 1970_1.R1040

5'-most EST LIB3028-050-Q1-B1-G11

Method BLASTX
NCBI GI g3386614
BLAST score 359
E value 3.0e-34
Match length 88
% identity 41

NCBI Description (AC004665) putative transcription factor SF3 [Arabidopsis

thaliana]

Seq. No. 2868

Contig ID 1972_1.R1040 5'-most EST gsv701044495.h1

Seq. No. 2869

Contig ID 1972_2.R1040

5'-most EST uC-gmropic103c02b1

Seq. No. 2870

Contig ID 1973_1.R1040

5'-most EST jC-gmro02910022c04a1

Method BLASTX
NCBI GI g3329368
BLAST score 251
E value 1.0e-21
Match length 61
% identity 79

NCBI Description (AF031244) nodulin-like protein [Arabidopsis thaliana]

Seq. No. 2871

Contig ID 1973_2.R1040 5'-most EST zhf700954080.h1

Seq. No. 2872

Contig ID 1973_4.R1040

5'-most EST jC-gmle01810077d06d1



```
2873
Seq. No.
                  1975_1.R1040
Contig ID
                  uC-gmrominsoy302h05b1
5'-most EST
                  {\tt BLASTX}
Method
                  g4262233
NCBI GI
                   460
BLAST score
E value
                   1.0e-45
                  246
Match length
% identity
                   46
NCBI Description (AC006200) hypothetical protein [Arabidopsis thaliana]
                   2874
Seq. No.
                   1976 1.R1040
Contig ID
                   jC-gmro02910046b07a1
5'-most EST
                   BLASTX
Method
                   g4567281
NCBI GI
                   201
BLAST score
                   3.0e-15
E value
Match length
                   114
                   43
% identity
NCBI Description (AC006841) unknown protein [Arabidopsis thaliana]
Seq. No.
                   1976 3.R1040
Contig ID
                   leu701150855.h1
5'-most EST
                   2876
Seq. No.
                   1977 1.R1040
Contig ID
                   LIB3028-050-Q1-B1-G9
5'-most EST
                   2877
Seq. No.
                   1979 1.R1040
Contig ID
                   LIB3028-026-Q1-B1-H2
5'-most EST
                   BLASTX
Method
                   q4204466
NCBI GI
                   582
BLAST score
                   8.0e-60
E value
                   222
Match length
                   56
% identity
                   (AF067417) mannose lectin [Dolichos lablab]
NCBI Description
                   2878
Seq. No.
                   1980 1.R1040
Contig ID
                   epx701107594.h1
 5'-most EST
                   BLASTN
Method
                    q3551246
NCBI GI
BLAST score
                    181
E value
                    6.0e-97
                    499
Match length
 % identity
                    85
                   Daucus carota mRNA for 181, partial cds
 NCBI Description
                    2879
 Seq. No.
```

1983 1.R1040

BLASTX

jC-gmle01810018c12a2

Contig ID

Method

5'-most EST



g2911799 NCBI GI 1303 BLAST score 1.0e-144 E value 328 Match length 75 % identity (AF008184) 4-coumarate:CoA ligase 1 [Populus balsamifera NCBI Description subsp. trichocarpa X Populus deltoides] 2880 Seq. No. 1984 1.R1040 Contig ID zhf700956267.h1 5'-most EST BLASTX Method g2262110 NCBI GI 235 BLAST score 7.0e-20 E value 65 Match length % identity (AC002343) zinc finger protein isolog [Arabidopsis NCBI Description thaliana] 2881 Seq. No. 1986 1.R1040 Contig ID asn701131155.hl 5'-most EST BLASTN Method g2970050 NCBI GI 380 BLAST score 0.0e + 00E value 692 Match length 89 % identity NCBI Description Vigna radiata mRNA for ARG10, complete cds Seq. No. 1986 2.R1040 Contig ID pxt700944387.hl 5'-most EST Seq. No. 2883 1986 3.R1040 Contig ID zzp700833378.hl 5'-most EST BLASTN Method g2970050 NCBI GI BLAST score 320 1.0e-180 E value Match length 727 % identity 88 Vigna radiata mRNA for ARG10, complete cds NCBI Description 2884 Seq. No. 1986 4.R1040

Contig ID 1986_4.R1040 5'-most EST fua701040205.h1

Method BLASTN
NCBI GI g2970050
BLAST score 215
E value 1.0e-117
Match length 351
% identity 90

NCBI Description Vigna radiata mRNA for ARG10, complete cds



```
2885
Seq. No.
                   1986 9.R1040
Contig ID
5'-most EST
                   jC-qmf102220097h01a1
                   BLASTN
Method
                   g2970050
NCBI GI
BLAST score
                   78
                   9.0e-36
E value
Match length
                   355
% identity
                   87
                   Vigna radiata mRNA for ARG10, complete cds
NCBI Description
                   2886
Seq. No.
                   1987 1.R1040
Contig ID
5'-most EST
                   g5753538
Method
                   BLASTX
NCBI GI
                   g2967456
                   473
BLAST score
                   2.0e-47
E value
Match length
                   164
% identity
                   54
                  (AB012048) sulfate transporter [Arabidopsis thaliana]
NCBI Description
                   2887
Seq. No.
                   1988 1.R1040
Contig ID
5'-most EST
                   LIB3028-040-Q1-B1-G1
                   2888
Seq. No.
                   1988 2.R1040
Contig ID
5'-most EST
                   zsg701118956.hl
                   2889
Seq. No.
                   1989 1.R1040
Contig ID
                   jC-gmro02910024d07a1
5'-most EST
                   BLASTX
Method
                   q3193330
NCBI GI
BLAST score
                   510
E value
                   2.0e-51
                   208
Match length
% identity
                   55
                   (AF069299) contains similarity to Medicago sativa corC
NCBI Description
                   (GB:L22305) [Arabidopsis thaliana]
Seq. No.
                   2890
Contig ID
                   1989 2.R1040
5'-most EST
                   6HA - \overline{0}1 - Q1 - B1 - G3
                   2891
Seq. No.
```

Contig ID 1989 4.R1040 5'-most EST zhf700962925.h1 BLASTX

Method g3738316 NCBI GI BLAST score 156 E value 1.0e-10 Match length 42 74 % identity

(AC005170) unknown protein [Arabidopsis thaliana] NCBI Description



Seq. No. 2892

Contig ID 1991 1.R1040 5'-most EST zzp700830841.h1

Method BLASTX
NCBI GI g3790581
BLAST score 237
E value 3.0e-19
Match length 157
% identity 37

NCBI Description (AF079179) RING-H2 finger protein RHB1a [Arabidopsis

thaliana]

Seq. No. 2893

Contig ID 1991 2.R1040

5'-most EST LIB3051-102-Q1-K1-C7

Method BLASTX
NCBI GI g3790581
BLAST score 214
E value 6.0e-17
Match length 111
% identity 38

NCBI Description (AF079179) RING-H2 finger protein RHB1a [Arabidopsis

thaliana]

Seq. No. 2894

Contig ID 1991_3.R1040 5'-most EST sat701009081.h1

Seq. No. 2895

Contig ID 1993_1.R1040

5'-most EST LIB3039-048-Q1-E1-D8

Method BLASTX
NCBI GI g4512671
BLAST score 381
E value 3.0e-36
Match length 162
% identity 57

NCBI Description (AC006931) unknown protein [Arabidopsis thaliana]

Seq. No. 2896

Contig ID 1993_2.R1040

5'-most EST LIB3039-006-Q1-E1-H10

Seq. No. 2897

Contig ID 1993_3.R1040 5'-most EST asn701138996.h1

Method BLASTX
NCBI GI 94512671
BLAST score 333
E value 1.0e-30
Match length 146
% identity 47

NCBI Description (AC006931) unknown protein [Arabidopsis thaliana]

Seq. No. 2898

Contig ID 1994 1.R1040 5'-most EST leu701150696.h1



Method BLASTN
NCBI GI g1658196
BLAST score 475
E value 0.0e+00
Match length 1107
% identity 86

NCBI Description Ricinus communis calreticulin mRNA, complete cds

Seq. No.

2899

Contig ID 5'-most EST 1994_2.R1040 LIB3087-001-Q1-K1-A11

Method BLASTX
NCBI GI g4467153
BLAST score 550
E value 3.0e-56
Match length 126

% identity 78

NCBI Description (AL035540) putative thaumatin-like protein [Arabidopsis

thaliana]

Seq. No.

2900

Contig ID

1997 1.R1040

5'-most EST

LIB3040-042-Q1-E1-A8

Method BLASTN
NCBI GI g2465009
BLAST score 98
E value 1.0e-47
Match length 170
% identity 89

NCBI Description Fragaria vesca mRNA for putative acyl carrier protein

Seq. No.

2901

Contig ID 1997_2.R1040
5'-most EST g5607168
Method BLASTN
NCBI GI g1838960
BLAST score 91
E value 3.0e-43
Match length 215

Match length 215 % identity 86

NCBI Description C.glauca mRNA for acyl carrier protein

Seq. No.

2902

Contig ID

1998_1.R1040

5'-most EST

uC-gmronoir037d06b1

Seq. No.

2903

Contig ID 5'-most EST

1998_2.R1040 zzp700832963.h1

Seq. No.

2904

Contig ID

1999_1.R1040

5'-most EST

LIB3028-050-Q1-B1-F1

Method BLASTX
NCBI GI g4249386
BLAST score 166
E value 1.0e-11



Match length 37 % identity 86

NCBI Description (AC005966) Strong similarity to gb_AF061286 gamma-adaptin 1 from Arabidopsis thaliana. EST gb_H37393 comes from this

gene. [Arabidopsis thaliana]

Seq. No. 2905

Contig ID 2002_1.R1040

5'-most EST LIB3028-050-Q1-B1-A8

Method BLASTX
NCBI GI g3928089
BLAST score 670
E value 4.0e-70
Match length 188
% identity 66

NCBI Description (AC005770) putative osr40 [Arabidopsis thaliana]

Seq. No. 2906

Contig ID 2002 2.R1040

5'-most EST LIB3049-014-Q1-E1-H9

Method BLASTX
NCBI GI g3928089
BLAST score 390
E value 2.0e-37
Match length 101
% identity 68

NCBI Description (AC005770) putative osr40 [Arabidopsis thaliana]

Seq. No. 2907

Contig ID 2003 1.R1040

5'-most EST LIB3139-094-P1-N1-B6

Method BLASTX
NCBI GI g2191150
BLAST score 852
E value 1.0e-149
Match length 355
% identity 45

NCBI Description (AF007269) similar to mitochondrial carrier family

[Arabidopsis thaliana]

Seq. No. 2908

Contig ID 2003 2.R1040

5'-most EST LIB3170-073-Q1-K1-G9

Method BLASTX
NCBI GI g2191150
BLAST score 345
E value 2.0e-32
Match length 81
% identity 37

NCBI Description (AF007269) similar to mitochondrial carrier family

[Arabidopsis thaliana]

Seq. No. 2909

Contig ID 2003_3.R1040

5'-most EST LIB3106-090-Q1-K1-C6

Method BLASTX NCBI GI g2191150



```
BLAST score 263
E value 8.0e-23
Match length 91
% identity 67
NCBI Description (AF007269) similar to mitochondrial carrier family [Arabidopsis thaliana]

Seq. No. 2910
Contig ID 2004_1.R1040
```

5'-most EST LIB3030-009-Q1-B1-H9
Method BLASTX
NCBI GI g728867
BLAST score 736
E value 1.0e-77
Match length 325

E value 1.06
Match length 325
% identity 44

NCBI Description ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR >gi 99694 pir S21961 proline-rich protein APG -

Arabidopsis thaliana >gi_22599_emb_CAA42925_ (X60377) APG

[Arabidopsis thaliana]

Seq. No. 2911

Contig ID 2005 1.R1040 5'-most EST vwf700674514.h1 Method BLASTX

NCBI GI g2583108
BLAST score 218
E value 2.0e-17
Match length 79
% identity 57

NCBI Description (AC002387) putative surface protein [Arabidopsis thaliana]

Seq. No. 2912

Contig ID 2005 2.R1040

5'-most EST LIB3106-004-Q1-K1-E8

Method BLASTX
NCBI GI g2507093
BLAST score 340
E value 3.0e-31
Match length 154
% identity 2

NCBI Description PERIOD CLOCK PROTEIN >gi_1196649 (M12039) ORF starts at 87, first start codon is found at 210.; putative [Mus musculus]

Seq. No. 2913

Contig ID 2005 3.R1040

5'-most EST LIB3094-048-Q1-K1-H6

Method BLASTX
NCBI GI g2583108
BLAST score 765
E value 2.0e-81
Match length 233
% identity 65

NCBI Description (AC002387) putative surface protein [Arabidopsis thaliana]

Seq. No. 2914

Contig ID 2006_1.R1040



```
5'-most EST
                  eep700870025.h1
                  BLASTX
Method
                  g1707074
NCBI GI
BLAST score
                  316
E value
                  1.0e-28
Match length
                  243
                  33
% identity
NCBI Description (U80450) M01E11.2 [Caenorhabditis elegans]
Seq. No.
                  2915
Contig ID
                  2010 1.R1040
5'-most EST
                  LIB3107-072-Q1-K1-C4
Method
                  BLASTX
NCBI GI
                  q4510421
BLAST score
                  196
                  1.0e-14
E value
Match length
                  264
                  27
% identity
NCBI Description (AC006929) unknown protein [Arabidopsis thaliana]
Seq. No.
                  2916
                  2011 1.R1040
Contig ID
5'-most EST
                  LIB3028-050-Q1-B1-C10
Seq. No.
                  2917
Contig ID
                  2012 1.R1040
5'-most EST
                  LIB3028-050-Q1-B1-C11
Method
                  BLASTX
NCBI GI
                  g1888357
BLAST score
                   444
E value
                   4.0e-44
Match length
                  141
% identity
                   58
                  (X98130) alpha-mannosidase [Arabidopsis thaliana]
NCBI Description
                  >gi_1890154_emb_CAA72432_ (Y11767) alpha-mannosidase
                  precursor [Arabidopsis thaliana]
Seq. No.
                   2918
Contig ID
                   2014 1.R1040
5'-most EST
                   jC-gmf102220142c10a1
Method
                   BLASTX
NCBI GI
                   g2842480
BLAST score
                   1456
                  1.0e-162
E value
                  359
Match length
                  79
% identity
                  (ALO21749) ADP, ATP carrier-like protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  2919
Contig ID
                   2014 2.R1040
```

5'-most EST jC-gmle01810061d02a1

Method BLASTX NCBI GI q2842480 BLAST score 548 5.0e-56 E value Match length 132



% identity 82
NCBI Description (AL021749) ADP, ATP carrier-like protein [Arabidopsis thaliana]

Seq. No. 2920

Contig ID 2014_3.R1040

5'-most EST uC-gmrominsoy129b01b1

Method BLASTX
NCBI GI g2842480
BLAST score 502
E value 8.0e-51
Match length 134
% identity 50

NCBI Description (AL021749) ADP, ATP carrier-like protein [Arabidopsis

thaliana]

Seq. No. 2921

2014 5.R1040 Contig ID 5'-most EST g5688431 Method BLASTN g1297067 NCBI GI BLAST score 58 E value 1.0e-23 Match length 170 % identity 84

NCBI Description T.turgidum ant gene (1494bp)

Seq. No. 2922

Contig ID 2016 1.R1040

5'-most EST LIB3139-095-P1-N1-G6

Method BLASTX
NCBI GI g2828280
BLAST score 651
E value 4.0e-68
Match length 184
% identity 70

NCBI Description (AL021687) putative protein [Arabidopsis thaliana]

>gi_2832633_emb_CAA16762_ (AL021711) putative protein

[Arabidopsis thaliana]

Seq. No. 2923

Contig ID 2016_3.R1040 5'-most EST zsg701119357.h1

Method BLASTX
NCBI GI g2828280
BLAST score 208
E value 1.0e-16
Match length 48
% identity 83

NCBI Description (AL021687) putative protein [Arabidopsis thaliana]

>gi_2832633_emb_CAA16762_ (AL021711) putative protein

[Arabidopsis thaliana]

Seq. No. 2924

Contig ID 2017_1.R1040

5'-most EST LIB3106-022-Q1-K1-E10

Method BLASTX



NCBI GI g3953473
BLAST score 504
E value 9.0e-51
Match length 140
% identity 71
NCBI Description (AC002328) F2202.18 [Arabidopsis thaliana]
Seq. No. 2925

Contig ID 2017_3.R1040 5'-most EST LIB3106-043-Q1-K1-F4 Method BLASTX NCBI GI g131194

NCBI GI g131194
BLAST score 387
E value 3.0e-37
Match length 89
% identity 85

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT V PRECURSOR

(PHOTOSYSTEM I 9 KD PROTEIN) (PSI-G) >gi_72686_pir__F1SP5

photosystem I chain V precursor - spinach

>gi_21299_emb_CAA31524_ (X13134) PSI subunit V preprotein (AA -69 to 98) [Spinacia oleracea] >gi_226167_prf__1413236B

photosystem I reaction center V [Spinacia oleracea]

Seq. No. 2926

Contig ID 2017 4.R1040

5'-most EST LIB3138-010-Q1-N1-H6

Seq. No. 2927

Contig ID 2019 1.R1040

5'-most EST LIB3028-050-Q1-B1-C8

Method BLASTX
NCBI GI g1076398
BLAST score 186
E value 8.0e-14
Match length 48
% identity 73

NCBI Description ribosomal protein L2 - Arabidopsis thaliana

>gi_572523_emb_CAA57902_ (X82556) ribosomal protein L2

[Arabidopsis thaliana]

Seq. No. 2928

Contig ID 2021 1.R1040

5'-most EST LIB3051-012-Q1-E1-H12

Method BLASTX
NCBI GI g2498973
BLAST score 288
E value 3.0e-25
Match length 270
% identity 31

NCBI Description SURFEIT LOCUS PROTEIN 1 >gi_1362943_pir__S57749 SURF-1

protein - human >gi_895849 emb_CAA84476 (Z35093) SURF-1 [Homo sapiens] >gi_4507319 ref_NP_003163.1_pSURF1 surfeit

Seq. No. 2929

Contig ID 2022 1.R1040

5'-most EST LIB3028-049-Q1-B1-H6



```
Seq. No.
Contig ID
```

2930

2024 1.R1040

5'-most EST jC-gmst02400048g07a1

Seq. No.

2931

Contig ID 5'-most EST 2024 2.R1040 jsh701066682.h1

Seq. No.

2932

Contig ID

2025 1.R1040

5'-most EST

uC-gmrominsoy271f10b1 BLASTX

Method NCBI GI BLAST score

q1362093 675

E value

1.0e-70

Match length % identity

212 63

NCBI Description

hypothetical protein (clone TPP15) - tomato (fragment)

>gi 924632 (U20595) unknown [Solanum lycopersicum]

Seq. No.

2933

Contig ID

2025 2.R1040

5'-most EST

LIB3106-112-Q1-K1-G9

Method NCBI GI BLASTX g1362093

BLAST score E value

222 1.0e-17

Match length

100 64

% identity NCBI Description

hypothetical protein (clone TPP15) - tomato (fragment)

>gi 924632 (U20595) unknown [Solanum lycopersicum]

Seq. No.

2934

Contig ID

2025 3.R1040 awf700842920.h1

5'-most EST Method

BLASTN q924631

NCBI GI BLAST score

59 2.0e-24

E value Match length

119

% identity

87

NCBI Description

Solanum lycopersicum predominantly pistil-, sepal-, and

fruit-expressed unknown protein mRNA, partial cds

Seq. No.

2935

Contig ID

2025 4.R1040

5'-most EST

uC-gmrominsoy243b12b1

Method NCBI GI BLASTX g1362093

BLAST score

636

E value

3.0e-66

Match length % identity

163

72

NCBI Description

hypothetical protein (clone TPP15) - tomato (fragment)

>gi_924632 (U20595) unknown [Solanum lycopersicum]

Seq. No.

2936



```
Contig ID 2025_5.R1040
5'-most EST LIB3106-049-Q1-K1-G2
Method BLASTN
NCBI GI g924631
BLAST score 51
E value 2.0e-19
```

119 86

% identity 86
NCBI Description Solanum lycopersicum predominantly pistil-, sepal-, and

fruit-expressed unknown protein mRNA, partial cds

Seq. No. 2937

Match length

Contig ID 2025 7.R1040

5'-most EST LIB3\overline{1}39-085-P4-N4-F12

Seq. No. 2938

Contig ID 2028_2.R1040 5'-most EST jex700909111.h1

Method BLASTX
NCBI GI g1834288
BLAST score 149
E value 2.0e-09
Match length 149
% identity 28

NCBI Description (AB000475) S.pombe TFA2 homolog [Schizosaccharomyces pombe]

Seq. No. 2939

Contig ID 2030 1.R1040

5'-most EST jC-gmf102220073c06a1

Method BLASTX
NCBI GI g2832625
BLAST score 337
E value 4.0e-31
Match length 194
% identity 38

NCBI Description (AL021711) putative protein [Arabidopsis thaliana]

Seq. No. 2940

Contig ID 2030 2.R1040

5'-most EST LIB3039-032-Q1-E1-D4

Method BLASTX
NCBI GI g2832625
BLAST score 605
E value 1.0e-62
Match length 257
% identity 45

NCBI Description (AL021711) putative protein [Arabidopsis thaliana]

Seq. No.

2941

Contig ID 2032_1.R1040

5'-most EST LIB3028-049-Q1-B1-H11

Seq. No. 2942

Contig ID 2034_1.R1040

5'-most EST LIB3106-060-Q1-K1-F3

Seq. No. 2943



2035 1.R1040 Contig ID 5'-most EST g5688044 BLASTN Method g2564046 NCBI GI 37 BLAST score 7.0e-11 E value Match length 269 86 % identity Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description MGI19, complete sequence [Arabidopsis thaliana] 2944 Seq. No. 2035 2.R1040 Contig ID LIB3028-010-Q1-B1-G3 5'-most EST BLASTN Method g2564046 NCBI GI 37 BLAST score 5.0e-11 E value 150 Match length 90 % identity Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone: NCBI Description MGI19, complete sequence [Arabidopsis thaliana] 2945 Seq. No. 2037 1.R1040 Contig ID LIB3051-104-Q1-K1-H3 5'-most EST BLASTX Method g1841355 NCBI GI 411 BLAST score 8.0e-40 E value 91 Match length 79 % identity (D85381) cytochrome c oxidase subunit Vb precursor [Oryza NCBI Description sativa] 2946 Seq. No. 2037 2.R1040 Contig ID 5'-most EST ncj700983001.h1 BLASTX Method g1841355 NCBI GI 173 BLAST score 3.0e-12 E value Match length 48 % identity 69 (D85381) cytochrome c oxidase subunit Vb precursor [Oryza NCBI Description sativa] 2947 Seq. No. 2038 1.R1040 Contig ID has700548129.h1 5'-most EST BLASTX Method g4063747 NCBI GI BLAST score 595

2.0e-61 E value 143 Match length 78 % identity

(AC005851) hypothetical protein [Arabidopsis thaliana] NCBI Description

Match length

% identity

189 83



```
2948
Seq. No.
                  2040 1.R1040
Contig ID
5'-most EST
                  LIB3028-020-Q1-B1-A1
Method
                  BLASTX
NCBI GI
                  q419963
BLAST score
                  285
E value
                  4.0e-25
Match length
                  85
% identity
                  59
                  snRNP protein B - fruit fly (Drosophila melanogaster)
NCBI Description
                  >gi 290272 (L02919) ribonucleoprotein [Drosophila
                  melanogaster]
                  2949
Seq. No.
                  2040 2.R1040
Contig ID
5'-most EST
                  g4397638
Method
                  BLASTN
NCBI GI
                  g3510341
BLAST score
                  46
E value
                  1.0e-16
Match length
                  114
% identity
                  85
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MFC16, complete sequence [Arabidopsis thaliana]
Seq. No.
                  2950
Contig ID
                  2042 1.R1040
5'-most EST
                  LIB3039-014-Q1-E1-A3
Method
                  BLASTX
NCBI GI
                  g2565429
BLAST score
                  144
                  1.0e-08
E value
Match length
                  36
% identity
                  75
NCBI Description (AF027686) glycine-rich protein [Onobrychis viciifolia]
Seq. No.
                  2951
                  2042 2.R1040
Contig ID
5'-most EST
                  LIB3039-035-Q1-E1-H4
Method
                  BLASTX
NCBI GI
                  g2565429
                  150
BLAST score
E value
                  2.0e-09
Match length
                  61
% identity
                  49
                  (AF027686) glycine-rich protein [Onobrychis viciifolia]
NCBI Description
                   2952
Seq. No.
Contig ID
                  2043 1.R1040
5'-most EST
                  uC-gmropic027h11b1
Method
                  BLASTX
NCBI GI
                  g1706130
BLAST score
                  830
                  6.0e-89
E value
```



NCBI Description CDPK-RELATED PROTEIN KINASE (PK421) >gi_2129917_pir__S60052

calcium-dependent protein kinase homolog - carrot

>gi_1103386_emb_CAA58750_ (X83869) CDPK-related protein

kinase [Daucus carota]

Seq. No. 2953

Contig ID 2043_2.R1040

5'-most EST fC-gmst700895016d3

Method BLASTX
NCBI GI g1706130
BLAST score 228
E value 1.0e-18
Match length 49
% identity 86

NCBI Description CDPK-RELATED PROTEIN KINASE (PK421) >gi 2129917 pir S60052

calcium-dependent protein kinase homolog - carrot

>gi_1103386_emb_CAA58750_ (X83869) CDPK-related protein

kinase [Daucus carota]

Seq. No. 2954

Contig ID 2044 1.R1040

5'-most EST jC-gmf102220137h06a1

Seq. No. 2955

Contig ID 2044 2.R1040

5'-most EST uC-gmropic100g07b1

Seq. No. 2956

Contig ID 2045 1.R1040

5'-most EST LIB3028-049-Q1-B1-G9

Method BLASTX
NCBI GI g2651310
BLAST score 434
E value 7.0e-43
Match length 111
% identity 74

NCBI Description (AC002336) putative PTR2-B peptide transporter [Arabidopsis

thaliana]

Seq. No. 2957

Contig ID 2046_1.R1040 5'-most EST gsv701052103.h1

Method BLASTX
NCBI GI g4455280
BLAST score 468
E value 1.0e-46
Match length 106
% identity 70

NCBI Description (AL035527) putative protein [Arabidopsis thaliana]

Seq. No. 2958

Contig ID 2047 1.R1040

5'-most EST jC-gmst02400041g03a1

Method BLASTX
NCBI GI g2760325
BLAST score 192
E value 2.0e-14



118 Match length % identity 44

NCBI Description (ACO02130) F1N21.10 [Arabidopsis thaliana]

Seq. No.

2959

Contig ID

2047 2.R1040

5'-most EST Method

LIB3170-047-Q1-J1-A2

NCBI GI BLAST score BLASTX q2648692 160

E value Match length 2.0e-10 262

% identity

28

NCBI Description (AE000975) NADH oxidase (noxA-5) [Archaeoglobus fulgidus]

Seq. No.

2960

Contig ID 5'-most EST

2047 3.R1040 pcp700995722.h1

Seq. No.

2961

Contig ID

2047 4.R1040

5'-most EST

jC-gmro02910067h12a1

Seq. No.

2962

Contig ID

2048 1.R1040

5'-most EST

LIB3028-049-Q1-B1-E5

Method BLASTX NCBI GI g3377517 BLAST score 346 2.0e-32 E value Match length 149

46 % identity

NCBI Description (AF073361) nitrate transporter NTL1 [Arabidopsis thaliana]

Seq. No.

2963

Contig ID

2050 1.R1040

5'-most EST

LIB3028-049-Q1-B1-E8

Seq. No.

2964

Contig ID

2051 1.R1040 dpv701101734.h1

5'-most EST Method

BLASTX

NCBI GI

g4262236

BLAST score

830

E value Match length 4.0e-92

% identity

256

NCBI Description

(AC006200) putative ribose 5-phosphate isomerase

[Arabidopsis thaliana]

Seq. No.

2965

Contig ID

2051 2.R1040

5'-most EST

jC-gmle01810073a01a1

Method BLASTX NCBI GI q4262236 BLAST score 362 3.0e-34 E value



Match length 106 % identity 73

NCBI Description (AC006200) putative ribose 5-phosphate isomerase

[Arabidopsis thaliana]

Seq. No. 2966

Contig ID 2053_1.R1040

5'-most EST LIB3072-021-Q1-E1-B12

Seq. No. 2967

Contig ID 2053_2.R1040

5'-most EST LIB3050-003-Q1-E1-B7

Seq. No. 2968

Contig ID 2054_1.R1040

5'-most EST uC-gmrominsoy244e12b1

Method BLASTN
NCBI GI g4159704
BLAST score 50
E value 8.0e-19
Match length 320

% identity 88
NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:

MCB17, complete sequence [Arabidopsis thaliana]

Seq. No. 2969

Contig ID 2054_2.R1040

5'-most EST LIB3028-049-Q1-B1-F11

Method BLASTN
NCBI GI g4159704
BLAST score 50
E value 6.0e-19
Match length 187
% identity 89

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:

MCB17, complete sequence [Arabidopsis thaliana]

Seq. No. 2970

Contig ID 2054_3.R1040 5'-most EST leu701152121.h1

Seq. No. 2971

Contig ID 2055 1.R1040

5'-most EST LIB3092-041-Q1-K1-G3

Method BLASTX
NCBI GI g1854443
BLAST score 1598
E value 1.0e-179
Match length 376
% identity 79

NCBI Description (D83970) CPRD8 protein [Vigna unguiculata]

Seq. No. 2972

Contig ID 2055_2.R1040

5'-most EST LIB3051-068-Q1-K1-A4

Method BLASTX NCBI GI g2232254

```
BLAST score
                   8.0e-57
E value
                   125
Match length
                   82
% identity
                   (AF005237) old-yellow-enzyme homolog [Catharanthus roseus]
NCBI Description
                    2973
Seq. No.
                   2055 4.R1040
Contig ID
                   uC-gmrominsoy113c06b1
5'-most EST
Method
                    BLASTX
                    q1854443
NCBI GI
                    433
BLAST score
                    6.0e-48
E value
Match length
                    111
                    85
% identity
                   (D83970) CPRD8 protein [Vigna unguiculata]
NCBI Description
Seq. No.
                    2974
                    2056 1.R1040
Contig ID
                    LIB3051-018-Q1-E1-G2
5'-most EST
                    2975
Seq. No.
                    2058 1.R1040
Contig ID
                    LIB3<del>1</del>06-010-Q1-K1-D7
5'-most EST
                    BLASTX
Method
                    g1710424
NCBI GI
                    455
BLAST score
                    5.0e-45
E value
                    199
Match length
                    52
% identity
                    50S RIBOSOMAL PROTEIN L21, CHLOROPLAST PRECURSOR (CL21)
NCBI Description
                    >gi 2129718 pir S71282 ribosomal protein L21 - Arabidopsis
                    tha\overline{\text{liana}} > g\overline{\text{i}} 11\overline{49}573 emb CAA89887 (Z49787) chloroplast
                    ribosomal large subunit protein L21 [Arabidopsis thaliana]
                    2976
Seq. No.
                    2059 1.R1040
Contig ID
```

 Seq. No.
 2976

 Contig ID
 2059 1.R1040

 5'-most EST
 k11701214615.h1

 Method
 BLASTX

 NCBI GI
 g2982463

NCBI GI g298246.
BLAST score 843
E value 1.0e-90
Match length 209
% identity 78

NCBI Description (AL022223) putative protein [Arabidopsis thaliana]

Seq. No. 2977

Contig ID 2059_2.R1040

5'-most EST LIB3028-018-Q1-B1-H10

Method BLASTX
NCBI GI g2982463
BLAST score 237
E value 8.0e-20
Match length 86
% identity 58

NCBI Description (AL022223) putative protein [Arabidopsis thaliana]



```
Seq. No.
                   2059 3.R1040
Contig ID
                   txt700734209.h1
5'-most EST
                   BLASTX
Method
                   g2982463
NCBI GI
BLAST score
                   146
                   2.0e-09
E value
                   37
Match length
% identity
                  (AL022223) putative protein [Arabidopsis thaliana]
NCBI Description
                   2979
Seq. No.
                   2060 1.R1040
Contig ID
                   LIB3109-014-Q1-K1-B11
5'-most EST
                   BLASTX
Method
                   q4580397
NCBI GI
BLAST score
                   498
                   3.0e-50
E value
Match length
                   204
                   50
% identity
NCBI Description (AC007171) putative RNA helicase [Arabidopsis thaliana]
                   2980
Seq. No.
                   2061 1.R1040
Contig ID
5'-most EST
                   zhf700962048.h1
                   2981
Seq. No.
                   2061 3.R1040
Contig ID
                   leu7\overline{0}1156542.h1
5'-most EST
                   BLASTN
Method
                   q1924955
NCBI GI
                   57
BLAST score
E value
                   2.0e-23
                   57
Match length
                   100
% identity
NCBI Description L.luteus 5S rRNA gene
                   2982
Seq. No.
                   2061 5.R1040
Contig ID
                   zhf700954372.h1
5'-most EST
Seq. No.
                   2983
                   2061 6.R1040
Contig ID
                   jC-gmst02400040c11a1
5'-most EST
                   2984
Seq. No.
                   2063 1.R1040
Contig ID
                   fC-gmf1700905086f4
5'-most EST
                   BLASTX
Method
                   g2598587
NCBI GI
```

2091 BLAST score 0.0e + 00E value 464 Match length 80 % identity

(Y15366) cycloartenol synthase [Medicago truncatula] NCBI Description

2985 Seq. No.



```
2065 1.R1040
Contig ID
                  LIB3028-048-Q1-B1-F7
5'-most EST
                   2986
Seq. No.
                   2068 1.R1040
Contig ID
                  LIB3028-049-Q1-B1-D7
5'-most EST
                   2987
Seq. No.
                   2069 1.R1040
Contig ID
                   sat701009654.hl
5'-most EST
                   2988
Seq. No.
                   2070 1.R1040
Contig ID
                   g5510085
5'-most EST
                   2989
Seq. No.
                   2070 2.R1040
Contig ID
                   g4396899
5'-most EST
                   {\tt BLASTX}
Method
                   g2583123
NCBI GI
                   525
BLAST score
                   2.0e-53
E value
                   145
Match length
                   68
% identity
                   (AC002387) putative nucleotide sugar epimerase [Arabidopsis
NCBI Description
                   thaliana]
                   2990
Seq. No.
                   2073 1.R1040
Contig ID
5'-most EST
                   LIB3028-039-Q1-B2-C1
                   2991
Seq. No.
                   2074 1.R1040
Contig ID
                   LIB3028-049-Q1-B1-B4
5'-most EST
                   BLASTX
Method
                   g4468979
NCBI GI
                   590
BLAST score
E value
                    9.0e-61
                    243
Match length
                    55
 % identity
NCBI Description (AL035605) putative protein [Arabidopsis thaliana]
 Seq. No.
                    2992
                    2074 2.R1040
 Contig ID
                    fua701041905.hl
 5'-most EST
                    BLASTX
 Method
                    q4468979
 NCBI GI
```

BLAST score 416 E value 7.0e-41 86 Match length % identity

(AL035605) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No.

2993

2075 1.R1040 Contig ID

LIB3028-049-Q1-B1-B5 5'-most EST



Seq. No. 2076 1.R1040 Contig ID vzy700751843.h1 5'-most EST BLASTX Method g3136057 NCBI GI 199 BLAST score 3.0e-15 E value 138 Match length 39 % identity (AL023592) actin-like protein [Schizosaccharomyces pombe] NCBI Description Seq. No. 2077 1.R1040 Contig ID LIB3028-049-Q1-B1-B8 5'-most EST BLASTN Method g1431738 NCBI GI 56 BLAST score 2.0e-22 E value 168 Match length 88 % identity Soybean (Glycine max) low MW heat shock protein gene NCBI Description (Gmhsp17.5-M) 2996 Seq. No. 2078 1.R1040 Contig ID LIB3028-049-Q1-B1-C10 5'-most EST BLASTX Method q4204793 NCBI GI BLAST score 693 3.0e-73 E value 152 Match length 43 % identity (U52079) P-glycoprotein [Solanum tuberosum] NCBI Description 2997 Seq. No. 2079 1.R1040 Contig ID 5'-most EST LIB3028-049-Q1-B1-C11 2998 Seq. No. 2080 1.R1040 Contig ID 5'-most EST hrw701062141.hl Method BLASTX NCBI GI g2894599 800 BLAST score 3.0e-85 E value 229 Match length % identity 69 (AL021889) putative protein [Arabidopsis thaliana] NCBI Description

2999 Seq. No.

2080 2.R1040 Contig ID bth700846937.h1 5'-most EST

BLASTXMethod g2894599 NCBI GI 822 BLAST score 8.0e-88 E value 225 Match length

% identity (AL021889) putative protein [Arabidopsis thaliana] NCBI Description 3000 Seq. No. 2080 3.R1040 Contig ID LIB3051-086-Q1-K1-A10 5'-most EST BLASTX Method g2894599 NCBI GI 539 BLAST score 4.0e-55 E value 120 Match length 80 % identity (AL021889) putative protein [Arabidopsis thaliana] NCBI Description 3001 Seq. No. 2080 4.R1040 Contig ID g4396557 BLASTX g2894599 295 1.0e-26

5'-most EST Method NCBI GI BLAST score E value 63 Match length % identity

(AL021889) putative protein [Arabidopsis thaliana] NCBI Description

3002 Seq. No.

2080 9.R1040 Contig ID jex700909422.hl 5'-most EST

BLASTX Method g2894599 NCBI GI 298 BLAST score E value 4.0e-27 65 Match length % identity 78

(AL021889) putative protein [Arabidopsis thaliana] NCBI Description

3003 Seq. No.

Contig ID 2080 10.R1040 $zsq7\overline{0}1127602.h1$ 5'-most EST

BLASTX Method q2894599 NCBI GI BLAST score 279 E value 4.0e-25 Match length 58 % identity

(AL021889) putative protein [Arabidopsis thaliana] NCBI Description

3004 Seq. No.

% identity

2082 1.R1040 Contig ID 5'-most EST leu701151978.hl

Method BLASTX g3819099 NCBI GI BLAST score 1152 E value 1.0e-126 264 Match length 77

(AJ009825) copper amine oxidase [Cicer arietinum] NCBI Description



```
Seq. No.
                   3005
Contig ID
                  2083 1.R1040
5'-most EST
                  hyd700724953.h1
                  BLASTX
Method
NCBI GI
                  g3292814
BLAST score
                  1095
                  1.0e-120
E value
Match length
                  317
                  70
% identity
NCBI Description (AL031018) putative protein [Arabidopsis thaliana]
Seq. No.
                  3006
                  2083 2.R1040
Contig ID
5'-most EST
                  LIB3138-059-Q1-N1-C8
Method
                  BLASTX
NCBI GI
                  g3292814
BLAST score
                   452
E value
                   6.0e-45
Match length
                  128
                   70
% identity
NCBI Description (AL031018) putative protein [Arabidopsis thaliana]
Seq. No.
                  2083 3.R1040
Contig ID
                  LIB3028-049-Q1-B1-B3
5'-most EST
                   3008
Seq. No.
Contig ID
                  2083 4.R1040
5'-most EST
                  g5606500
Method
                  BLASTX
                  g3292814
NCBI GI
BLAST score
                   316
                   4.0e-29
E value
Match length
                  94
                   68
% identity
NCBI Description
                  (AL031018) putative protein [Arabidopsis thaliana]
Seq. No.
                   3009
                   2085 1.R1040
Contig ID
5'-most EST
                   k11701203522.h2
Method
                  BLASTX
                   g4457221
NCBI GI
BLAST score
                   372
                   5.0e-35
E value
                  161
Match length
                  56
% identity
                  (AF127797) putative bZIP DNA-binding protein [Capsicum
NCBI Description
                  chinensel
Seq. No.
                   3010
Contig ID
                   2085 2.R1040
5'-most EST
                  LIB3087-011-Q1-K1-C12
Method
                   BLASTX
```

BLAST score 361 E value 7.0e-34

g4457221

NCBI GI



```
Match length
                   98
                   77
% identity
NCBI Description
                   (AF127797) putative bZIP DNA-binding protein [Capsicum
                  chinensel
                  3011
Seq. No.
                  2085 3.R1040
Contig ID
5'-most EST
                  LIB3049-006-Q1-E1-A5
Method
                  BLASTX
NCBI GI
                   q4457221
BLAST score
                   385
E value
                   1.0e-36
Match length
                  101
% identity
                   77
                  (AF127797) putative bZIP DNA-binding protein [Capsicum
NCBI Description
                  chinense]
                  3012
Seq. No.
                   2085 4.R1040
Contig ID
5'-most EST
                  LIB3093-008-Q1-K1-F4
                   3013
Seq. No.
Contig ID
                  2085 5.R1040
5'-most EST
                  jC-gmst02400047b01d1
Seq. No.
                   3014
Contig ID
                  2087 1.R1040
5'-most EST
                   jC-gmro02800040d02a1
                  BLASTX
Method
NCBI GI
                  q2190542
BLAST score
                  573
E value
                   5.0e-59
                  211
Match length
% identity
                  54
NCBI Description (AC001229) F5I14.5 gene product [Arabidopsis thaliana]
                  3015
Seq. No.
                  2089_1.R1040
Contig ID
5'-most EST
                  vzy700750788.h1
Method
                  BLASTX
NCBI GI
                  q3158476
BLAST score
                  1232
E value
                  1.0e-136
Match length
                  285
% identity
                  82
NCBI Description (AF067185) aquaporin 2 [Samanea saman]
                  3016
Seq. No.
                  2089 2.R1040
Contig ID
5'-most EST
                  LIB3050-014-Q1-E1-H4
                  BLASTX
Method
```

NCBI GI g3158476 BLAST score 628 E value 1.0e-65 Match length 145 % identity 83

NCBI Description (AF067185) aquaporin 2 [Samanea saman]



```
3017
Seq. No.
                   2094 1.R1040
Contig ID
                  LIB3093-001-Q1-K1-E6
5'-most EST
                  BLASTX
Method
                   g3608137
NCBI GI
                   660
BLAST score
E value
                   1.0e-68
Match length
                   333
                   43
% identity
                  (AC005314) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3018
Seq. No.
                   2094 2.R1040
Contig ID
                   jC-gmro02910037h07d1
5'-most EST
                   3019
Seq. No.
                   2095 1.R1040
Contig ID
                   sat701009411.h1
5'-most EST
Method
                   BLASTX
                   g4263722
NCBI GI
BLAST score
                   1724
                   0.0e + 00
E value
Match length
                   635
                   62
% identity
                   (AC006223) putative glucan synthase [Arabidopsis thaliana]
NCBI Description
                   3020
Seq. No.
                   2096 1.R1040
Contig ID
                   fC-gmro700865146f1
5'-most EST
                   BLASTN
Method
                   g3641864
NCBI GI
BLAST score
                   260
                   1.0e-144
E value
                   697
Match length
                   84
% identity
                   Cicer arietinum mRNA for beta-galactosidase, clone
NCBI Description
                   CanBGal-4
                   3021
Seq. No.
                   2096 2.R1040
Contig ID
                   zhf700955904.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3126969
                   729
BLAST score
                    4.0e-77
E value
                   223
Match length
% identity
                   (AF061808) chalcone isomerase [Elaeagnus umbellata]
NCBI Description
                    3022
 Seq. No.
                    2096 3.R1040
 Contig ID
```

5'-most EST uC-gmflminsoy003e05b1

Method BLASTX
NCBI GI g3641863
BLAST score 1913
E value 0.0e+00



```
496
Match length
                   75
% identity
```

NCBI Description (AJ005042) beta-galactosidase [Cicer arietinum]

Seq. No.

3023

Contig ID 5'-most EST 2096 4.R1040

Method

uC-gmflminsoy001a11b1 BLASTX

NCBI GI BLAST score E value

q3860420 1813 0.0e + 00

Match length % identity

383

NCBI Description (AJ011047) exo galactanase [Lupinus angustifolius]

Seq. No.

3024

Contig ID

2096 5.R1040

5'-most EST

 $\verb"uC-gmflminsoy034h12b1"$

Seq. No.

3025

Contig ID

2096 6.R1040

5'-most EST

LIB3072-052-Q1-E1-D10

Method NCBI GI

BLASTN g3860419

BLAST score

65

E value

6.0e-28

Match length % identity

117 90

NCBI Description Lupinus angustifolius mRNA for exo galactanase

Seq. No.

3026

Contig ID 5'-most EST

2096 8.R1040 sat701007604.h1

Method

BLASTN

NCBI GI

g3860419

BLAST score

119

E value

3.0e-60

Match length % identity

199 90

NCBI Description Lupinus angustifolius mRNA for exo galactanase

Seq. No.

3027

Contig ID

2097 1.R1040

5'-most EST

LIB3028-048-Q1-B1-G12

Method NCBI GI BLASTX g2760326

BLAST score

1178

E value

Match length

1.0e-129

% identity

531

NCBI Description (AC002130) F1N21.11 [Arabidopsis thaliana]

Seq. No.

3028

46

Contig ID 5'-most EST 2097 2.R1040

g5676937 BLASTX

Method NCBI GI

g2760326



```
BLAST score
                   5.0e-12
E value
                  73
Match length
                   55
% identity
NCBI Description (AC002130) F1N21.11 [Arabidopsis thaliana]
Seq. No.
                   3029
                  2097 4.R1040
Contig ID
                  zhf700953819.h1
5'-most EST
                  BLASTX
Method
                  g2760326
NCBI GI
BLAST score
                  176
                   1.0e-12
E value
                   100
Match length
                   40
% identity
                  (AC002130) F1N21.11 [Arabidopsis thaliana]
NCBI Description
                   3030
Seq. No.
                   2098 1.R1040
Contig ID
                   fC-gmse700674243f2
5'-most EST
                   BLASTX
Method
                   q4056500
NCBI GI
BLAST score
                   165
                   3.0e-11
E value
                   52
Match length
                   60
% identity
                   (AC005896) putative acetyltransferase [Arabidopsis
NCBI Description
                   thaliana]
                   3031
Seq. No.
                   2098 2.R1040
Contig ID
                   LIB3028-048-Q1-B1-G3
5'-most EST
Method
                   BLASTX
                   q1653293
NCBI GI
BLAST score
                   175
E value
                   1.0e-22
                   96
Match length
% identity
                   59
                  (D90912) ornithine acetyltransferase [Synechocystis sp.]
NCBI Description
                   3032
Seq. No.
                   2099 1.R1040
Contig ID
5'-most EST
                   vwf700675854.h1
Seq. No.
                   3033
                   2100 1.R1040
Contig ID
                   zhf700957066.hl
5'-most EST
Method
                   BLASTN
                   g516248
NCBI GI
BLAST score
                   146
                   4.0e-76
E value
                   310
Match length
% identity
                   87
                   A.thaliana gene for porphobilinogen deaminase
NCBI Description
```

Seq. No. 3034

Contig ID 2100_2.R1040



```
jC-gmro02800036a04d1
5'-most EST
                  BLASTN
Method
                  g516248
NCBI GI
                  97
BLAST score
                  4.0e-47
E value
                  177
Match length
                  89
% identity
NCBI Description A.thaliana gene for porphobilinogen deaminase
Seq. No.
                  2102 1.R1040
Contig ID
                  LIB3049-042-Q1-E1-A8
5'-most EST
                   3036
Seq. No.
                   2102 2.R1040
Contig ID
                  LIB3049-004-Q1-E1-G5
5'-most EST
                   3037
Seq. No.
                   2103 1.R1040
Contig ID
                  LIB3093-029-Q1-K1-E5
5'-most EST
                   BLASTN
Method
                   g3928149
NCBI GI
                   412
BLAST score
                   0.0e + 00
E value
                   852
Match length
                   87
% identity
NCBI Description Cicer arietinum mRNA for hypothetical protein
                   3038
Seq. No.
                   2103 2.R1040
Contig ID
                   uC-qmropic028g12b1
5'-most EST
                   BLASTX
Method
                   q3928150
NCBI GI
                   230
BLAST score
                   2.0e-40
E value
                   124
Match length
% identity
                   (AJ131049) hypothetical protein [Cicer arietinum]
NCBI Description
Seq. No.
                   2103 3.R1040
Contig ID
                   kl1701203849.hl
5'-most EST
Method
                   BLASTN
                   q3928149
NCBI GI
BLAST score
                   69
E value
                   1.0e-30
Match length
                   197
                   84
 % identity
NCBI Description Cicer arietinum mRNA for hypothetical protein
                   3040
 Seq. No.
                   2105 1.R1040
 Contig ID
                   txt700735241.h1
 5'-most EST
```

BLASTX Method g3641837 NCBI GI

1027 BLAST score 0.0e+00E value



Match length 550 % identity 86

NCBI Description (AL023094) Nonclathrin coat protein gamma - like protein

[Arabidopsis thaliana]

Seq. No. 3041

Contig ID 2105_4.R1040 5'-most EST hrw701058021.h1

Method BLASTX
NCBI GI g3641837
BLAST score 262
E value 6.0e-23
Match length 55
% identity 91

NCBI Description (AL023094) Nonclathrin coat protein gamma - like protein

[Arabidopsis thaliana]

Seq. No. 3042

Contig ID 2106_1.R1040

5'-most EST LIB3109-014-Q1-K1-A11

Method BLASTX
NCBI GI g4539001
BLAST score 1288
E value 1.0e-142
Match length 318
% identity 78

NCBI Description (AL049481) putative protein [Arabidopsis thaliana]

Seq. No. 3043

Contig ID 2109_1.R1040 5'-most EST zhf700961869.h1

Method BLASTX
NCBI GI g2160165
BLAST score 220
E value 9.0e-18
Match length 69
% identity 59

NCBI Description (AC000132) F21M12.11 gene product [Arabidopsis thaliana]

Seq. No. 3044

Contig ID 2110_1.R1040

5'-most EST uC-gmropic080d08b1

Method BLASTX
NCBI GI g2894598
BLAST score 625
E value 1.0e-102
Match length 335
% identity 63

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 3045

Contig ID 2111_1.R1040 5'-most EST epx701109255.h1

Seq. No. 3046

Contig ID 2112 1.R1040 5'-most EST zzp700833245.h1



BLASTX Method g3377850 NCBI GI 325 BLAST score 9.0e-30 E value 133 Match length 72 % identity

(AF076274) contains simlarity to Canis familiaris signal NCBI Description peptidase complex 25 kDa subunit (GB:U12687) [Arabidopsis

thaliana]

3047 Seq. No.

2113 1.R1040 Contig ID

LIB3109-038-Q1-K1-B10 5'-most EST

BLASTX Method g549010 NCBI GI 1992 BLAST score 0.0e + 00E value 433 Match length 89 % identity

EUKARYOTIC PEPTIDE CHAIN RELEASE FACTOR SUBUNIT 1 (ERF1) NCBI Description

(OMNIPOTENT SUPPRESSOR PROTEIN 1 HOMOLOG) (SUP1 HOMOLOG) >gi_322554_pir__S31328 omnipotent suppressor protein SUP1

homolog (clone G18) - Arabidopsis thaliana

>gi 16514_emb_CAA49172_ (X69375) similar to yeast

omnipotent suppressor protein SUP1 (SUP45) [Arabidopsis thaliana] >gi_1402882_emb_CAA66813_ (X98130) eukaryotic early release factor subunit 1-like protein [Arabidopsis

thaliana] >gi_1495249 emb_CAA66118 (X97486) eRF1-3

[Arabidopsis Thaliana]

3048 Seq. No.

2113 3.R1040 Contig ID trc700563986.h1 5'-most EST

3049 Seq. No.

2116 1.R1040 Contiq ID

LIB3139-098-P1-N1-G1 5'-most EST

3050 Seq. No.

2119 1.R1040 Contig ID

uC-qmflminsoy063c12b1 5'-most EST

Method BLASTX NCBI GI q2435522 BLAST score 1333 1.0e-147 E value 388 Match length % identity

(AF024504) contains similarity to other AMP-binding enzymes NCBI Description

[Arabidopsis thaliana]

3051 Seq. No.

2119 2.R1040 Contig ID

uC-gmropic010b07b1 5'-most EST

Method BLASTX q1903034 NCBI GI 195 BLAST score 1.0e-14 E value



```
Match length
                  94
                   43
% identity
                  (X94625) amp-binding protein [Brassica napus]
NCBI Description
                  3052
Seq. No.
                  2120 1.R1040
Contig ID
5'-most EST
                  jC-gmst02400060f04a1
                  3053
Seq. No.
                  2120 2.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy028d12b1
Method
                  BLASTX
NCBI GI
                  q3063710
BLAST score
                  208
E value
                  3.0e-16
Match length
                  150
% identity
                  37
NCBI Description (AL022537) putative protein [Arabidopsis thaliana]
Seq. No.
Contig ID
                  2123 1.R1040
5'-most EST
                  zsq701123264.h1
Method
                  BLASTX
NCBI GI
                  g3319340
BLAST score
                  231
E value
                  5.0e-19
Match length
                  91
% identity
                  54
                  (AF077407) contains similarity to E. coli cation transport
NCBI Description
                  protein ChaC (GB:D90756) [Arabidopsis thaliana]
                   3055
Seq. No.
Contig ID
                  2125 1.R1040
5'-most EST
                  uC-gmflminsoy078e04b1
Seq. No.
                   3056
                   2126 1.R1040
Contig ID
5'-most EST
                  LIB3028-048-Q1-B1-D9
Method
                  BLASTX
NCBI GI
                  g3513738
BLAST score
                  149
                  2.0e-09
E value
                  37
Match length
                   68
% identity
                  (AF080118) similar to the GDSL family of lipolytic enzymes
NCBI Description
                   [Arabidopsis thaliana]
                   3057
Seq. No.
                  2131 1.R1040
Contig ID
5'-most EST
                  LIB3087-009-Q1-K1-E1
Method
                  BLASTX
```

NCBI GI g2618721 BLAST score 716 2.0e-75 E value Match length 231 67 % identity

NCBI Description (U49072) IAA16 [Arabidopsis thaliana]



```
3058
Seq. No.
                   2131 2.R1040
Contig ID
                   LIB3039-034-Q1-E1-F1
5'-most EST
                   BLASTX
Method
                   g114734
NCBI GI
BLAST score
                   655
                   2.0e-68
E value
                   226
Match length
                   65
% identity
                  AUXIN-INDUCED PROTEIN AUX28 >gi_81759_pir__A28993 aux28
NCBI Description
                   protein - soybean >gi_169921 (J\overline{0}3919) auxin-regulated
                   protein (Aux28) [Glycine max]
                   3059
Seq. No.
                   2131 3.R1040
Contig ID
                   LIB3107-066-Q1-K1-D4
5'-most EST
                   BLASTX
Method
                   g2618721
NCBI GI
BLAST score
                   660
                   6.0e-69
E value
                   234
Match length
                   61
% identity
                   (U49072) IAA16 [Arabidopsis thaliana]
NCBI Description
                   3060
Seq. No.
                   2131 4.R1040
Contig ID
                   sat7\overline{0}1006155.h1
5'-most EST
                   BLASTX
Method
                   g2618721
NCBI GI
BLAST score
                   116
                   1.0e-08
E value
                   69
Match length
% identity
                   (U49072) IAA16 [Arabidopsis thaliana]
NCBI Description
                   3061
Seq. No.
Contig ID
                   2132 1.R1040
                   LIB3028-048-Q1-B1-C8
5'-most EST
                   BLASTX
Method
                   g2352492
NCBI GI
BLAST score
                   488
E value
                   6.0e-49
Match length
                   115
                   78
% identity
                   (AF005047) transport inhibitor response 1 [Arabidopsis
NCBI Description
                   thaliana] >gi 2352494 (AF005048) transport inhibitor
                   response 1 [Arabidopsis thaliana]
                   3062
Seq. No.
                   2132 2.R1040
Contig ID
                   uC-gmrominsoy125g04b1
5'-most EST
```

Method BLASTX
NCBI GI g2352492
BLAST score 329
E value 2.0e-30
Match length 83



% identity

NCBI Description

(AF005047) transport inhibitor response 1 [Arabidopsis thaliana] >gi_2352494 (AF005048) transport inhibitor

response 1 [Arabidopsis thaliana]

Seq. No.

Contig ID

3063 2134 1.R1040

5'-most EST

LIB3139-039-P1-N1-H4

Method NCBI GI BLASTX q2947062

BLAST score E value

550 8.0e-56

Match length % identity

534 35

NCBI Description

(AC002521) unknown protein [Arabidopsis thaliana]

Seq. No.

3064

Contig ID

2134 2.R1040

5'-most EST

uC-gmropic083d02b1

Seq. No.

Contig ID 5'-most EST 3065

2134 3.R1040 pmv700894820.hl

Seq. No.

3066

Contig ID

2135 1.R1040

5'-most EST

LIB3167-023-P4-K4-B8

Method NCBI GI BLAST score BLASTX g2495180 1073

E value

1.0e-117

Match length

251 84

% identity NCBI Description

PORPHOBILINOGEN DEAMINASE PRECURSOR (PBG)

(HYDROXYMETHYLBILANE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN

SYNTHASE) >gi_421882_pir__S35873 hydroxymethylbilane

synthase (EC 4.3.1.8) - garden pea >gi_541971_pir__JQ2278 hydroxymethylbilane synthase (EC 4.3.1.8) precursor garden pea chloroplast >gi_313724_emb_CAA51820_ (X73418)

hydroxymethylbilane synthase [Pisum sativum]

Seq. No.

3067

Contig ID

2135 2.R1040

5'-most EST

LIB3028-048-Q1-B1-D10

Method NCBI GI BLASTN q313723

BLAST score

134

E value

3.0e-69

Match length

322

% identity

85

NCBI Description P.sativum hemC mRNA for hydroxymethylbilane synthase

Seq. No.

3068

Contig ID 5'-most EST 2135 5.R1040 qsv701054125.hl

Method

BLASTX

NCBI GI g2495179



BLAST score 268 E value 1.0e-37 Match length 146 % identity 79

PORPHOBILINOGEN DEAMINASE PRECURSOR (PBG)

(HYDROXYMETHYLBILANE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN SYNTHASE) >gi_1084340_pir__S50762 hydroxymethylbilane

synthase (EC 4.3.1.8) precursor - Arabidopsis thaliana >gi_313150 emb_CAA51941_ (X73535) hydroxymethylbilane synthase [Arabidopsis thaliana] >gi_313838 emb_CAA52061_

 $(\bar{X}73839)$ hydroxymethylbilane synthase [Arabidopsis

thaliana]

3069

Seq. No.

Contig ID 5'-most EST

NCBI Description

2135_6.R1040 ncj700975534.h1

Method BLASTX
NCBI GI g2495180
BLAST score 271
E value 5.0e-36
Match length 89
% identity 90

NCBI Description PORPHOBILINOGEN DEAMINASE PRECURSOR (PBG)

(HYDROXYMETHYLBILANE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN

SYNTHASE) >gi_421882_pir__S35873 hydroxymethylbilane synthase (EC 4.3.1.8) - garden pea >gi_541971_pir_JQ2278 hydroxymethylbilane synthase (EC 4.3.1.8) precursor - garden pea chloroplast >gi_313724_emb_CAA51820_ (X73418)

hydroxymethylbilane synthase [Pisum sativum]

Seq. No. 3070

Contig ID 2136_1.R1040

5'-most EST LIB3138-022-Q1-N1-D5

Method BLASTN
NCBI GI g1381085
BLAST score 322
E value 0.0e+00
Match length 578
% identity 89

NCBI Description Pisum sativum imidazoleglycerol-phosphate dehydratase

(PSHIS3) mRNA, complete cds

Seq. No. 3071

Contig ID 2137_1.R1040

5'-most EST LIB3028-048-Q1-B1-D12

Seq. No.

3072

Contig ID 2138_1.R1040

5'-most EST LIB3093-027-Q1-K1-B1

Method BLASTN
NCBI GI g3885514
BLAST score 178
E value 3.0e-95
Match length 334
% identity 88

NCBI Description Medicago sativa clone MS56 unknown mRNA



```
3073
Seq. No.
                   2138 2.R1040
Contig ID
                  LIB3040-036-Q1-E1-A7
5'-most EST
                  BLASTN
Method
                   g3885514
NCBI GI
BLAST score
                   189
                   1.0e-102
E value
Match length
                   369
% identity
                   88
                  Medicago sativa clone MS56 unknown mRNA
NCBI Description
                   3074
Seq. No.
                   2138 4.R1040
Contig ID
                   asn7\overline{0}1132221.h2
5'-most EST
                   BLASTN
Method
                   g3885514
NCBI GI
BLAST score
                   86
                   2.0e-40
E value
                   186
Match length
                   87
% identity
NCBI Description Medicago sativa clone MS56 unknown mRNA
                   3075
Seq. No.
                   2138 5.R1040
Contig ID
                   LIB3040-018-Q1-E1-H12
5'-most EST
                   BLASTN
Method
                   g3885514
NCBI GI
                   163
BLAST score
                   2.0e-86
E value
                   311
Match length
                   88
% identity
                   Medicago sativa clone MS56 unknown mRNA
NCBI Description
                   3076
Seq. No.
                   2139 1.R1040
Contig ID
                   LIB3028-048-Q1-B1-D3
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4371290
                   512
BLAST score
                   1.0e-51
E value
                   190
Match length
                   67
% identity
                   (AC006260) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3077
Seq. No.
                   2139 2.R1040
Contig ID
5'-most EST
                   LIB3049-051-Q1-E1-B8
                   BLASTX
Method
                   g4371290
NCBI GI
BLAST score
                   374
                   8.0e-36
E value
Match length
                   89
% identity
                   (AC006260) unknown protein [Arabidopsis thaliana]
NCBI Description
```

3078

2140 1.R1040

Seq. No.

Contig ID



```
awf700843415.h1
5'-most EST
                  BLASTX
Method
                  g464987
NCBI GI
BLAST score
                  776
                  1.0e-82
E value
                  149
Match length
% identity
                  97
                  UBIQUITIN-CONJUGATING ENZYME E2-17 KD 10 (UBIQUITIN-PROTEIN
NCBI Description
                  LIGASE 10) (UBIQUITIN CARRIER PROTEIN 10)
                  >gi 421858_pir__S32672 ubiquitin--protein ligase (EC
                  6.3.2.19) UBC10 - Arabidopsis thaliana
                  >gi 297878 emb_CAA78715_ (Z14991) ubiquitin conjugating
                  enzyme [Arabidopsis thaliana] >gi_349213 (L00640) ubiquitin
                  conjugating enzyme [Arabidopsis thaliana]
                  3079
Seq. No.
                  2140 2.R1040
Contig ID
                  LIB3050-003-Q1-E1-F12
5'-most EST
                  BLASTN
Method
NCBI GI
                   g297877
                  179
BLAST score
                   1.0e-95
E value
Match length
                   443
                   85
% identity
                  A.thaliana UBC10 mRNA for ubiquitin conjugating enzyme
NCBI Description
                   homolog >gi_349212_gb_L00640_ATHUBCC Arabidopsis thaliana
                   ubiquitin conjugating enzyme mRNA, complete cds
                   3080
Seq. No.
                   2140 3.R1040
Contig ID
                   leu7\overline{0}1147536.h1
5'-most EST
                   BLASTN
Method
                   g297877
NCBI GI
                   175
BLAST score
                   3.0e-93
E value
                   447
Match length
                   85
% identity
                   A.thaliana UBC10 mRNA for ubiquitin conjugating enzyme
NCBI Description
                   homolog >gi_349212_gb_L00640_ATHUBCC Arabidopsis thaliana
                   ubiquitin conjugating enzyme mRNA, complete cds
                   3081
Seq. No.
                   2140 5.R1040
Contig ID
5'-most EST
                   jC-qmro02910066g10a1
Method
                   BLASTN
```

Method BLASTN
NCBI GI g388206
BLAST score 37
E value 2.0e-11
Match length 69

% identity 88

NCBI Description Lycopersicon esculentum ubiquitin carrier protein (Ubc) mRNA, complete cds. >gi 2086269 gb I38279 I38279 Sequence

109 from patent US

Seq. No. 3082

Contig ID 2140 6.R1040

5'-most EST LIB3092-003-Q1-K1-D3



BLASTN Method g297877 NCBI GI 102 BLAST score 5.0e-50 E value 290 Match length 84 % identity

A.thaliana UBC10 mRNA for ubiquitin conjugating enzyme NCBI Description homolog >gi_349212_gb_L00640_ATHUBCC Arabidopsis thaliana

ubiquitin conjugating enzyme mRNA, complete cds

3083 Seq. No.

2141 1.R1040 Contig ID $k117\overline{0}1212974.h1$ 5'-most EST

BLASTX Method g2464855 NCBI GI 551 BLAST score 2.0e-56 E value 176 Match length % identity

(Z99707) myb-related protein [Arabidopsis thaliana] NCBI Description

Seq. No.

2144 2.R1040 Contig ID

jC-gmf102220108a01a1 5'-most EST

3085 Seq. No.

2144 3.R1040 Contig ID

jC-gmf102220090d03a1 5'-most EST

BLASTN Method g3033512 NCBI GI BLAST score 557 0.0e + 00E value 1251 Match length 87 % identity

Phaseolus vulgaris rubisco activase (Rcal) mRNA, complete NCBI Description

3086 Seq. No.

2144 4.R1040 Contig ID

LIB3138-105-Q1-N1-F11 5'-most EST

BLASTN Method q4406529 NCBI GI BLAST score 499 0.0e + 00E value 1411 Match length 88 % identity

Vigna radiata rubisco activase (Rca) mRNA, chloroplast gene NCBI Description

encoding chloroplast protein, complete cds

3087 Seq. No.

2148 1.R1040 Contig ID

uC-gmrominsoy176b07b1 5'-most EST

Method BLASTX g4115925 NCBI GI 798 BLAST score 6.0e-85 E value 270 Match length



% identity NCBI Description

(AF118222) contains similarity to RNA recognition motifs (Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana] >gi_4539439_emb_CAB40027.1_ (AL049523) RNA-binding protein [Arabidopsis thaliana]

Seq. No. 3088

Contig ID 2148_2.R1040

5'-most EST jC-gmst02400014a06a1

Method BLASTX
NCBI GI g4115925
BLAST score 683
E value 9.0e-72
Match length 231

Match length 231 % identity 43

NCBI Description (AF118222) contains similarity to RNA recognition motifs

(Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana] >gi 4539439_emb_CAB40027.1_ (AL049523) RNA-binding protein

[Arabidopsis thaliana]

Seq. No. 3089

Contig ID 2150_1.R1040 5'-most EST pmv700894902.h1

Seq. No. 3090

Contig ID 2150_2.R1040 5'-most EST sat701013106.h1

Seq. No. 3091

Contig ID 2151_1.R1040 5'-most EST dpv701103482.h1

Method BLASTX
NCBI GI g2462927
BLAST score 178
E value 6.0e-13

Match length 64 % identity 52

NCBI Description (Y10342) putative amidase [Arabidopsis thaliana]

Seq. No. 3092

Contig ID 2152 1.R1040 5'-most EST kl1701205744.h1

Method BLASTX
NCBI GI g285741
BLAST score 694
E value 1.0e-72
Match length 376
% identity 45

NCBI Description (D14550) EDGP precursor [Daucus carota]

Seq. No. 3093

Contig ID 2152 2.R1040

5'-most EST jC-gmro02910034c10d1

Method BLASTX
NCBI GI g285741
BLAST score 320
E value 2.0e-29



104 Match length

% identity NCBI Description

(D14550) EDGP precursor [Daucus carota]

Seq. No. Contig ID 3094

2152 3.R1040

uC-gmropic116e09b1 5'-most EST

Seq. No.

3095

Contig ID 5'-most EST 2153 1.R1040 has700548121.h1

Method NCBI GI BLAST score E value

g4006816 166 4.0e-11

BLASTX

Match length % identity

99 43

NCBI Description

(AC005970) hypothetical protein, 3' partial [Arabidopsis

thaliana]

Seq. No.

3096

2153 2.R1040 Contig ID 5'-most EST ncj700984512.hl

Seq. No. Contig ID 3097

5'-most EST

2154 1.R1040

LIB3028-048-Q1-B1-B2

BLASTX Method g3201619 NCBI GI BLAST score 557 E value 4.0e-57 Match length 163

% identity

65

NCBI Description

(AC004669) dihydroxypolyprenylbenzoate methyltransferase

[Arabidopsis thaliana]

Seq. No.

3098

Contig ID

2155 1.R1040

5'-most EST

q4294625

Seq. No.

3099

Contig ID

2156 1.R1040

5'-most EST '

uC-gmrominsoy173a11b1

Method NCBI GI BLAST score BLASTX q2181180

E value

1149

Match length

1.0e-125

% identity

699 38

NCBI Description

(Z84377) xylosidase [Aspergillus niger]

Seq. No.

3100

Contig ID 5'-most EST 2158 1.R1040 zhf700955352.h1

Method BLASTX NCBI GI q4115547 282 BLAST score



```
7.0e-32
E value
                  155
Match length
                  51
% identity
                  (AB012634) lectin [Robinia pseudoacacia]
NCBI Description
Seq. No.
                  2161 1.R1040
Contig ID
                  jC-gmle01810044a01d1
5'-most EST
                  BLASTX
Method
                  g1388078
NCBI GI
                  307
BLAST score
                  1.0e-27
E value
                  104
Match length
                   47
% identity
NCBI Description (U35826) thioredoxin h [Arabidopsis thaliana]
                   3102
Seq. No.
                   2161 2.R1040
Contig ID
                  uC-gmrominsoy101a09b1
5'-most EST
                  BLASTX
Method
                   g1388078
NCBI GI
                   265
BLAST score
                   7.0e-23
E value
                   106
Match length
                   42
% identity
                  (U35826) thioredoxin h [Arabidopsis thaliana]
NCBI Description
                   3103
Seq. No.
                   2161 3.R1040
Contig ID
5'-most EST
                   zsg701119390.hl
                   BLASTX
Method
                   g129960
NCBI GI
BLAST score
                   494
                   2.0e-49
E value
                   282
Match length
                   42
% identity
NCBI Description
                   4-NITROPHENYLPHOSPHATASE (PNPPASE)
                   3104
Seq. No.
Contig ID
                   2161 4.R1040
5'-most EST
                   LIB3139-033-P1-N1-F1
Method
                   BLASTX
NCBI GI
                   q1388078
BLAST score
                   269
E value
                   1.0e-23
Match length
                   88
% identity
                   57
                   (U35826) thioredoxin h [Arabidopsis thaliana]
NCBI Description
                   3105
Seq. No.
```

Contig ID 2163_1.R1040
5'-most EST g4297823
Method BLASTX
NCBI GI g3372230
BLAST score 461
E value 9.0e-46
Match length 140



```
% identity
                   (AF017074) RNA polymerase I, II and III 16.5 kDa subunit
NCBI Description
                   [Arabidopsis thaliana]
                   3106
Seq. No.
                  2164 1.R1040
Contig ID
                  LIB3049-050-Q1-E1-B5
5'-most EST
Method
                  BLASTX
NCBI GI
                   q1041825
BLAST score
                   228
E value
                   4.0e-18
                   283
Match length
                   25
% identity
                   (U26175) LPG2 [Leishmania donovani]
NCBI Description
                   >gi_1583457_prf__2120442A LPG2 gene [Leishmania donovani]
                   3107
Seq. No.
                   2166 1.R1040
Contig ID
                   txt700735601.hl
5'-most EST
                   3108
Seq. No.
                   2167 1.R1040
Contig ID
                   LIB3028-047-Q1-B1-H1
5'-most EST
                   3109
Seq. No.
                   2168 1.R1040
Contig ID
                   LIB3170-057-Q1-K1-A4
5'-most EST
                   3110
Seq. No.
                   2168 2.R1040
Contig ID
5'-most EST
                   jC-gmf102220104a05a1
                   3111
Seq. No.
                   2168 3.R1040
Contig ID
5'-most EST
                   LIB3039-046-Q1-E1-E4
                   3112
Seq. No.
                   2169 1.R1040
Contig ID
                   jC-gmst02400026h10a1
5'-most EST
                   BLASTX
Method
                   q4115386
NCBI GI
BLAST score
                   317
E value
                   3.0e-29
Match length
                   120
% identity
                   (AC005967) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3113
Seq. No.
Contig ID
                   2169 3.R1040
5'-most EST
                   g4396207
Method
                   BLASTX
                   g4115386
NCBI GI
BLAST score
                   212
                   7.0e-17
E value
                   59
Match length
```

(AC005967) unknown protein [Arabidopsis thaliana]

76

% identity

NCBI Description



```
3114
Seq. No.
                  2171 1.R1040
Contig ID
                  LIB3109-048-Q1-K1-A2
5'-most EST
                  BLASTX
Method
                  g2288981
NCBI GI
BLAST score
                  512
                   7.0e-52
E value
Match length
                   148
                   67
% identity
                   (AC002335) calcium binding protein isolog [Arabidopsis
NCBI Description
                   thaliana] >gi 3763938 (AC004450) putative calcium binding
                   protein [Arabidopsis thaliana]
                   3115
Seq. No.
                   2171 2.R1040
Contig ID
                   dpv701099410.h1
5'-most EST
                   BLASTX
Method
                   g2288981
NCBI GI
BLAST score
                   527
                   3.0e-53
E value
                   148
Match length
% identity
                   69
                   (AC002335) calcium binding protein isolog [Arabidopsis
NCBI Description
                   thaliana] >gi_3763938 (AC004450) putative calcium binding
                   protein [Arabidopsis thaliana]
                   3116
Seq. No.
                   2171 3.R1040
Contig ID
                   rca701000403.hl
5'-most EST
                   3117
Seq. No.
                   2174 1.R1040
Contig ID
                   LIB3040-061-Q1-E1-H9
5'-most EST
                   BLASTX
Method
                   q4512699
NCBI GI
                   1164
BLAST score
                   1.0e-128
E value
                   300
Match length
                   75
 % identity
                    (AC006569) putative NADH-ubiquinone oxireductase
NCBI Description
                    [Arabidopsis thaliana]
                   3118
 Seq. No.
                   2175 1.R1040
 Contig ID
                   LIB3028-047-Q1-B1-H7
 5'-most EST
Method
                   BLASTX
```

NCBI GI g2088650
BLAST score 244

E value 2.0e-20

Match length 74 % identity 68

NCBI Description (AF002109) peroxisomal ATP/ADP carrier protein isolog

[Arabidopsis thaliana]

Seq. No. 3119

Contig ID 2176 1.R1040



```
LIB3028-040-Q1-B1-A11
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2281334
                  409
BLAST score
                  9.0e-40
E value
                  174
Match length
% identity
                  51
NCBI Description (U83619) putative pectate lyase [Arabidopsis thaliana]
                  3120
Seq. No.
                  2177 1.R1040
Contig ID
                  kl1701210975.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4544449
                   831
BLAST score
                   3.0e-89
E value
                   203
Match length
                   76
% identity
NCBI Description (AC006592) putative peroxidase [Arabidopsis thaliana]
Seq. No.
                   3121
                   2179 1.R1040
Contig ID
                   jC-qmro02910030d10a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q543867
                   1284
BLAST score
                   1.0e-142
E value
                   313
Match length
                   84
% identity
                   ATP SYNTHASE GAMMA CHAIN, MITOCHONDRIAL PRECURSOR
NCBI Description
                   >gi_1076684_pir__A47493 H+-transporting ATP synthase (EC
                   3.6.1.34) gamma chain precursor - sweet potato
                   >gi 303626_dbj_BAA03526_ (D14699) F1-ATPase gammma subunit
                   [Ipomoea batatas]
                   3122
Seq. No.
                   2179 2.R1040
Contig ID
5'-most EST
                   leu701152070.hl
                   BLASTN
Method
                   q397554
NCBI GI
                   159
BLAST score
                   7.0e-84
E value
Match length
                   287
 % identity
                   89
NCBI Description N.sylvestris psaH gene for photosystem I psaH protein
                   3123
 Seq. No.
                   2179 4.R1040
 Contig ID
                   LIB3051-045-Q1-K1-A10
 5'-most EST
                   BLASTX
Method
                   g543867
 NCBI GI
 BLAST score
                   433
 E value
                   1.0e-42
 Match length
                   114
 % identity
                   81
                   ATP SYNTHASE GAMMA CHAIN, MITOCHONDRIAL PRECURSOR
 NCBI Description
                   >gi_1076684_pir__A47493 H+-transporting ATP synthase (EC
```



3.6.1.34) gamma chain precursor - sweet potato >gi_303626_dbj_BAA03526_ (D14699) F1-ATPase gammma subunit [Ipomoea batatas]

Seq. No. 3124

Contig ID 2179 5.R1040

5'-most EST LIB3073-015-Q1-K1-C10

Method BLASTN
NCBI GI g397556
BLAST score 158
E value 2.0e-83
Match length 306
% identity 88

NCBI Description N.tabacum psaH gene for photosystem I psaH protein

Seq. No. 3125

Contig ID 2179_6.R1040

5'-most EST LIB3073-016-Q1-K1-B2

Method BLASTN
NCBI GI g397554
BLAST score 124
E value 5.0e-63
Match length 284
% identity 86

NCBI Description N.sylvestris psaH gene for photosystem I psaH protein

Seq. No. 3126

Contig ID 2179 7.R1040

5'-most EST LIB3074-004-Q1-K1-H1

Method BLASTN
NCBI GI g397554
BLAST score 85
E value 9.0e-40
Match length 249
% identity 85

NCBI Description N.sylvestris psaH gene for photosystem I psaH protein

Seq. No. 3127

Contig ID 2179_8.R1040 5'-most EST g5677856

Seq. No. 3128

Contig ID 2179_9.R1040

5'-most EST jC-qmro02910030a06d1

Method BLASTN
NCBI GI g286033
BLAST score 44
E value 2.0e-15
Match length 104
% identity 86

NCBI Description Ipomoea batatas (sweet potato) mRNA for F1-ATPase

gamma-subunit

Seq. No. 3129

Contig ID 2179_11.R1040 5'-most EST pmv700892366.h1

Method BLASTN

E value

Match length

291



```
q397556
NCBI GI
BLAST score
                  92
                  3.0e-44
E value
                  160
Match length
                  89
% identity
NCBI Description N.tabacum psaH gene for photosystem I psaH protein
Seq. No.
                  2179 13.R1040
Contiq ID
                   fua701042954.h1
5'-most EST
                   3131
Seq. No.
                   2183 1.R1040
Contig ID
                  LIB3139-091-P1-N1-C10
5'-most EST
                   BLASTX
Method
                   g2662375
NCBI GI
                   934
BLAST score
                   1.0e-101
E value
Match length
                   443
                   42
% identity
NCBI Description (D89060) oligosaccharyltransferase [Homo sapiens]
                   3132
Seq. No.
                   2184 1.R1040
Contig ID
                   uC-qmflminsoy069g07b1
5'-most EST
                   BLASTX
Method
                   g2505879
NCBI GI
                   252
BLAST score
E value
                   4.0e-29
                   268
Match length
                   37
% identity
                  (Y12776) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3133
Seq. No.
                   2185 1.R1040
Contig ID
                   LIB3028-030-Q1-B1-G10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4539301
BLAST score
                   252
                   2.0e-21
E value
                   142
Match length
                   39
% identity
                   (AL049480) putative mitochondrial protein [Arabidopsis
NCBI Description
                   thaliana]
                   3134
Seq. No.
                   2185 2.R1040
Contig ID
                   LIB3093-027-Q1-K1-A11
5'-most EST
                   3135
Seq. No.
                   2186 1.R1040
Contig ID
5'-most EST
                   g5606255
Method
                   BLASTX
NCBI GI
                   g4455223
                   1033
BLAST score
                   1.0e-112
```



% identity (AL035440) putative DNA binding protein [Arabidopsis NCBI Description thaliana] 3136 Seq. No. 2186 2.R1040 Contig ID 5'-most EST LIB3028-012-Q1-B1-C9 Method BLASTX g4455223 NCBI GI 651 BLAST score 8.0e-68 E value 180 Match length 45 % identity (AL035440) putative DNA binding protein [Arabidopsis NCBI Description thaliana] 3137 Seq. No. 2186 3.R1040 Contig ID pcp700991124.h1 5'-most EST BLASTX Method NCBI GI g4455223 BLAST score 424 4.0e-59 E value Match length 209 % identity 32 (AL035440) putative DNA binding protein [Arabidopsis NCBI Description thaliana] Seq. No. 3138 2186 4.R1040 Contig ID 5'-most EST LIB3167-023-P4-K4-G9 BLASTX Method q1899188 NCBI GI BLAST score 266 4.0e-23 E value Match length 72 % identity (U90212) DNA binding protein ACBF [Nicotiana tabacum] NCBI Description Seq. No. 2186 5.R1040 Contig ID 5'-most EST uC-gmropic068b09b1 Method BLASTX NCBI GI q4455223 BLAST score 331 E value 3.0e - 31Match length 100

28 % identity

(AL035440) putative DNA binding protein [Arabidopsis NCBI Description

thaliana]

3140 Seq. No.

Contig ID 2186 7.R1040 $uxk7\overline{0}0669935.h1$ 5'-most EST

BLASTX Method q4455223 NCBI GI BLAST score 229



```
5.0e-19
E value
Match length
                   63
% identity
                   37
NCBI Description
                   thaliana]
```

(AL035440) putative DNA binding protein [Arabidopsis

3141 Seq. No.

2187 1.R1040 Contig ID zpv700759026.h15'-most EST

Seq. No.

2190 1.R1040 Contig ID

jC-gmle01810035e07a2 5'-most EST

3142

3143 Seq. No.

2191 1.R1040 Contig ID

LIB3028-047-Q1-B1-F8 5'-most EST

Seq. No. 3144

2193 1.R1040 Contig ID

LIB3050-004-Q1-E1-E12 5'-most EST

Method BLASTX g4263712 NCBI GI 486 BLAST score E value 1.0e-48 119 Match length 76 % identity

(AC006223) putative ribosomal protein S12 [Arabidopsis NCBI Description

thaliana]

3145 Seq. No.

2193 2.R1040 Contig ID kl1701212032.h1 5'-most EST

BLASTX Method g4263712 NCBI GI BLAST score 488 6.0e-49E value 119 Match length % identity

(AC006223) putative ribosomal protein S12 [Arabidopsis NCBI Description

thaliana]

3146 Seq. No.

Contig ID 2193 3.R1040 leu701155992.h1 5'-most EST

BLASTX Method q4263712 NCBI GI BLAST score 356 E value 9.0e-34 84 Match length 79 % identity

(AC006223) putative ribosomal protein S12 [Arabidopsis NCBI Description

thaliana]

3147 Seq. No.

2193 4.R1040 Contig ID bth700848563.h1 5'-most EST



```
BLASTX
Method
                   q4263712
NCBI GI
BLAST score
                   373
                   4.0e-38
E value
                   107
Match length
                   74
% identity
                   (AC006223) putative ribosomal protein S12 [Arabidopsis
NCBI Description
                   thaliana]
                   3148
Seq. No.
                   2193 5.R1040
Contig ID
                   LIB3139-075-P1-N1-D5
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4263712
BLAST score
                   401
                   9.0e-39
E value
Match length
                   121
                   65
% identity
                   (AC006223) putative ribosomal protein S12 [Arabidopsis
NCBI Description
                   thaliana]
                   3149
Seq. No.
                   2193 7.R1040
Contig ID
                   LIB3040-009-Q1-E1-A6
5'-most EST
                   3150
Seq. No.
                   2194 1.R1040
Contig ID
                   fua701040372.h1
5'-most EST
                   BLASTX
Method
                   g2369714
NCBI GI
                   2838
BLAST score
                   0.0e+00
E value
                   865
Match length
                   89
% identity
NCBI Description (Z97178) elongation factor 2 [Beta vulgaris]
                   3151
Seq. No.
                   2194 2.R1040
Contig ID
5'-most EST
                   rlr700899150.h1
                   BLASTN
Method
                   q2369713
NCBI GI
BLAST score
                   54
                   2.0e-21
E value
Match length
                   162
 % identity
                   85
                   Beta vulgaris cDNA for elongation factor
NCBI Description
                   3152
 Seq. No.
                   2194 4.R1040
 Contig ID
                   LIB3093-010-Q1-K1-H2
 5'-most EST
                   BLASTX
 Method
 NCBI GI
                   g2369714
 BLAST score
                   208
 E value
                    3.0e-16
 Match length
                    55
```

(Z97178) elongation factor 2 [Beta vulgaris]

73

% identity

NCBI Description



Seq. No. 3153

Contig ID 2196 1.R1040

5'-most EST uC-gmrominsoy199h08b1

Method BLASTX
NCBI GI g2492519
BLAST score 2180
E value 0.0e+00
Match length 426
% identity 99

NCBI Description 26S PROTEASE REGULATORY SUBUNIT 7 (26S PROTEASOME SUBUNIT

7) >gi_1395191_dbj_BAA13021_ (D86121) 26S proteasome ATPase

subunit [Spinacia oleracea]

Seq. No. 3154

Contig ID 2196_2.R1040

5'-most EST jC-gmro02910005d08d1

Method BLASTN
NCBI GI g1395190
BLAST score 45
E value 5.0e-16
Match length 85
% identity 88

NCBI Description Spinacia oleracea L. mRNA for 26S proteasome ATPase

subunit, complete cds

Seq. No. 3155

Contig ID 2196_3.R1040 5'-most EST rca701001049.h1

Method BLASTX
NCBI GI g2492519
BLAST score 272
E value 3.0e-24
Match length 58
% identity 90

NCBI Description 26S PROTEASE REGULATORY SUBUNIT 7 (26S PROTEASOME SUBUNIT

7) >gi_1395191_dbj_BAA13021_ (D86121) 26S proteasome ATPase

subunit [Spinacia oleracea]

Seq. No. 3156

Contig ID 2196_4.R1040 5'-most EST zzp700831319.h1 Method BLASTN

Method BLASTN
NCBI GI g1395190
BLAST score 46
E value 7.0e-17
Match length 86
% identity 88

NCBI Description Spinacia oleracea L. mRNA for 26S proteasome ATPase

subunit, complete cds

Seq. No. 3157

Contig ID 2197_1.R1040 5'-most EST epx701110235.h1

Method BLASTN NCBI GI g1707656 BLAST score 65

```
5.0e-28
E value
                  105
Match length
                   90
% identity
NCBI Description P.sativum mRNA for DnaJ-like protein
                   3158
Seq. No.
                   2198 1.R1040
Contig ID
                  LIB3109-014-Q1-K1-D7
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4106538
                   1078
BLAST score
                   1.0e-118
E value
                   282
Match length
                   72
% identity
                   (AF104220) gamma-tocopherol methyltransferase [Arabidopsis
NCBI Description
                   thaliana]
                   3159
Seq. No.
                   2198 2.R1040
Contig ID
                   taw700655856.h1
5'-most EST
                   BLASTX
Method
                   g4106538
NCBI GI
                   244
BLAST score
                   1.0e-20
E value
                   60
Match length
                   75
% identity
                   (AF104220) gamma-tocopherol methyltransferase [Arabidopsis
NCBI Description
                   thaliana]
                   3160
Seq. No.
                   2200 1.R1040
Contig ID
                   vwf700678153.hl
5'-most EST
                   BLASTX
Method
                   g1350680
NCBI GI
                   1614
BLAST score
                   1.0e-180
E value
                   404
Match length
                   76
```

% identity

60S RIBOSOMAL PROTEIN L1 NCBI Description

3161 Seq. No.

2200 3.R1040 Contig ID gsv701051391.h1 5'-most EST

BLASTX Method a2459420 NCBI GI 716 BLAST score 1.0e-75 E value 140 Match length % identity

(AC002332) putative ribosomal protein L17 [Arabidopsis NCBI Description

thaliana]

3162 Seq. No.

2200 4.R1040 Contig ID

LIB3106-005-Q1-K1-D8 5'-most EST

Method BLASTN g310932 NCBI GI



BLAST score 161 E value 4.0e-85 Match length 421 % identity 85

NCBI Description Nicotiana tabacum ribosomal protein L17 mRNA, complete cds

Seq. No. 3163

Contig ID 2200_5.R1040

5'-most EST LIB3051-015-Q1-E1-G11

Method BLASTN
NCBI GI g310932
BLAST score 145
E value 1.0e-75
Match length 413
% identity 84

NCBI Description Nicotiana tabacum ribosomal protein L17 mRNA, complete cds

Seq. No. 3164

Contig ID 2200_8.R1040 5'-most EST kwa701015594.h1

Seq. No. 3165

Contig ID 2201_1.R1040

5'-most EST LIB3093-044-Q1-K1-G2

Method BLASTX
NCBI GI g3249071
BLAST score 560
E value 3.0e-57
Match length 164
% identity 60

NCBI Description (AC004473) Contains similarity to protein-tyrosine

phosphatase 2 gb_L15420 from Dictyostelium discoideum. EST

gb_N38718 comes from this g [Arabidopsis thaliana]

Seq. No. 3166

Contig ID 2201_2.R1040 5'-most EST asn701134505.h1

Method BLASTX
NCBI GI g3249071
BLAST score 155
E value 1.0e-10
Match length 41
% identity 68

NCBI Description (AC004473) Contains similarity to protein-tyrosine

phosphatase 2 gb_L15420 from Dictyostelium discoideum. EST

gb_N38718 comes from this g [Arabidopsis thaliana]

Seq. No. 3167

Contig ID 2204_1.R1040 5'-most EST ssr700561073.h1

Method BLASTN
NCBI GI g938299
BLAST score 690
E value 0.0e+00
Match length 1290
% identity 89

NCBI Description V.ungiculata mRNA for unknown protein (A3 gene)



```
3168
Seq. No.
                  2204 4.R1040
Contig ID
                  g5509857
5'-most EST
Seq. No.
                  3169
                  2204 6.R1040
Contig ID
5'-most EST
                  zsg701119048.hl
Method
                  BLASTN
NCBI GI
                  q938299
                  114
BLAST score
                  2.0e-57
E value
                  158
Match length
                  93
% identity
NCBI Description V.ungiculata mRNA for unknown protein (A3 gene)
                   3170
Seq. No.
                   2205 1.R1040
Contig ID
5'-most EST
                  LIB3028-047-Q1-B1-D8
                   3171
Seq. No.
                   2206 1.R1040
Contig ID
                   LIB3092-007-Q1-K1-C7
5'-most EST
                   BLASTN
Method
NCBI GI
                   g1370204
BLAST score
                   391
                   0.0e+00
E value
                   615
Match length
                   91
% identity
                   L.japonicus mRNA for small GTP-binding protein, RAN1B
NCBI Description
Seq. No.
                   2206 2.R1040
Contig ID
                   hrw701058039.h1
5'-most EST
                   BLASTN
Method
                   q1370204
NCBI GI
                   125
BLAST score
                   1.0e-63
E value
                   226
Match length
                   92
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAN1B
                   3173
Seq. No.
                   2206 3.R1040
Contig ID
                   smc700745908.h1
5'-most EST
                   BLASTN
Method
                   q1370202
NCBI GI
                   207
BLAST score
                   1.0e-112
E value
                   347
Match length
 % identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAN1A
```

Seq. No. 3174

Contig ID 2206_5.R1040 5'-most EST leu701145214.h1

Method BLASTX



```
q2149051
NCBI GI
BLAST score
                  631
                  1.0e-116
E value
                  222
Match length
% identity
                  93
                  (U73810) small Ras-like GTP-binding protein [Arabidopsis
NCBI Description
                  thaliana]
                   3175
Seq. No.
                   2206 10.R1040
Contig ID
                  pxt700944036.hl
5'-most EST
                  BLASTN
Method
                   g1370202
NCBI GI
                   149
BLAST score
E value
                   3.0e-78
                   209
Match length
% identity
                   93
                  L.japonicus mRNA for small GTP-binding protein, RAN1A
NCBI Description
Seq. No.
                   2206 11.R1040
Contig ID
                 hrw701060904.hl
5'-most EST
Method
                   BLASTN
NCBI GI
                   q1370202
                   112
BLAST score
                   4.0e-56
E value
                   234
Match length
                   94
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAN1A
                   3177
Seq. No.
                   2206 13.R1040
Contig ID
                   uxk700672208.h1
5'-most EST
                   BLASTN
Method
                   g1370204
NCBI GI
                   56
BLAST score
                   1.0e-22
E value
                   153
Match length
                   92
 % identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAN1B
                   3178
 Seq. No.
                   2208 1.R1040
 Contig ID
 5'-most EST
                   LIB3039-031-Q1-E1-A3
 Method
                   BLASTX
                   q4538897
 NCBI GI
                    539
 BLAST score
                    5.0e-55
 E value
                    154
 Match length
                    63
 % identity
                   (ALO49482) AX110P-like protein [Arabidopsis thaliana]
 NCBI Description
                    3179
 Seq. No.
 Contig ID
                    2210 1.R1040
                    sat701007262.h2
 5'-most EST
```

BLASTX

g2505876

Method NCBI GI



```
BLAST score
                   336
                   4.0e-31
E value
                  89
Match length
                   75
% identity
NCBI Description (Y12776) MYB-related protein [Arabidopsis thaliana]
                   3180
Seq. No.
                   2211 1.R1040
Contig ID
5'-most EST
                   leu701155294.hl
                   BLASTX
Method
NCBI GI
                   q3176684
BLAST score
                   365
                   1.0e-34
E value
                   148
Match length
% identity
                   51
                   (AC003671) Contains similarity to equilibratiave nucleoside
NCBI Description
                   transporter 1 gb_U81375 from Homo sapiens. ESTs gb_N65317,
                   gb_T20785, gb_AA586285 and gb_AA712578 come from this gene.
                   [Arabidopsis thaliana]
                   3181
Seq. No.
                   2213 1.R1040
Contig ID
                   gsv7\overline{0}1047793.h1
5'-most EST
Method
                   BLASTX
                   g464365
NCBI GI
BLAST score
                   1212
                   1.0e-133
E value
                   296
Match length
% identity
                   78
                   PEROXIDASE P7 >gi 66306 pir OPNB7 peroxidase (EC 1.11.1.7)
NCBI Description
                   - turnip
                   3182
Seq. No.
                   2214 1.R1040
 Contig ID
                   LIB3072-049-Q1-K1-B6
 5'-most EST
                   BLASTX
Method
                   q3608481
NCBI GI
                   623
BLAST score
                   7.0e-65
 E value
                   148
 Match length
                   79
 % identity
                   (AF088913) ribosomal protein L27a [Petunia x hybrida]
 NCBI Description
 Seq. No.
                   3183
                    2214 2.R1040
 Contig ID
                   LIB3093-012-Q1-K1-F8
 5'-most EST
                   BLASTX
 Method
                    q1710530
 NCBI GI
 BLAST score
                    455
                    3.0e-45
 E value
                    148
 Match length
 % identity
                    60S RIBOSOMAL PROTEIN L27A >gi_2129719_pir__S71256
 NCBI Description
                    ribosomal protein L27a - Arabidopsis thaliana
```

L27a [Arabidopsis thaliana]

>gi 1107487 emb_CAA63025_ (X91959) 60S ribosomal protein



```
      Seq. No.
      3184

      Contig ID
      2216_1.R1040

      5'-most EST
      pxt700943374.h1
```

Seq. No. 3185 Contig ID 2217_1.R1040 5'-most EST LIB3051-050-Q1-K1-A10

Method BLASTX
NCBI GI g4263722
BLAST score 766
E value 1.0e-81
Match length 203

% identity 70 NCBI Description (AC006223) putative glucan synthase [Arabidopsis thaliana]

Contig ID 2218_1.R1040
5'-most EST LIB3039-049-Q1-E1-D2
Method BLASTX
NCBI GI g4185505
BLAST score 266

BLAST score 266
E value 6.0e-23
Match length 91
% identity 54

Seq. No.

NCBI Description (AF101038) nonspecific lipid-transfer protein precursor

[Brassica napus]

3187 Seq. No. 2220 1.R1040 Contig ID 5'-most EST q5752955 BLASTX Method q549750 NCBI GI BLAST score 233 5.0e-19 E value 96 Match length % identity 48

NCBI Description HYPOTHETICAL 29.4 KD PROTEIN IN STE6-LOS1 INTERGENIC REGION >gi 539221_pir__S38045 hypothetical protein YKL207w - yeast

(Saccharomyces cerevisiae) >gi_486369_emb_CAA82052_(Z28207) ORF YKL207w [Saccharomyces cerevisiae]

Seq. No. 3188

Contig ID 2220_2.R1040

5'-most EST LIB3051-028-Q1-K1-G7

Seq. No. 3189

Contig ID 2220_3.R1040 5'-most EST kl1701209834.h1

Method BLASTX
NCBI GI g549750
BLAST score 218
E value 2.0e-17
Match length 147
% identity 37
NCBI Description HYPOTHET

HYPOTHETICAL 29.4 KD PROTEIN IN STE6-LOS1 INTERGENIC REGION >gi_539221_pir__S38045 hypothetical protein YKL207w - yeast

(Saccharomyces cerevisiae) >gi_486369_emb_CAA82052_



(Z28207) ORF YKL207w [Saccharomyces cerevisiae]

Seq. No. 3190 Contig ID 2220_4.R1040

Contig ID 2220_4.R1040 5'-most EST sat701004550.h1

Seq. No. 3191

Contig ID 2220_5.R1040

5'-most EST LIB3050-024-Q1-K1-D10

Seq. No. 3192

Contig ID 2221 1.R1040

5'-most EST LIB3028-049-Q1-B1-A12

Method BLASTX
NCBI GI g3142301
BLAST score 1220
E value 1.0e-134
Match length 497

% identity 53

NCBI Description (AC002411) Contains similarity to neural cell adhesion

molecule 2, large isoform precursor gb_M76710 from Xenopus laevis, and beta transducin from S. cerevisiae gb_Q05946. ESTs gb_N65081 gb_Z30910, gb_Z34190, gb_Z34611, gb_R30101,

gb H3630

Seq. No. 3193

Contig ID 2222_1.R1040 5'-most EST zsg701121747.h1

Method BLASTX
NCBI GI g2351374
BLAST score 821
E value 5.0e-88
Match length 180
% identity 90

NCBI Description (U54560) putative 26S proteasome subunit athMOV34

[Arabidopsis thaliana]

Seq. No. 3194

Contig ID 2224_1.R1040 5'-most EST crh700850891.h1

Method BLASTX
NCBI GI g3269287
BLAST score 751
E value 5.0e-80
Match length 177
% identity 78

NCBI Description (AL030978) GH3 like protein [Arabidopsis thaliana]

Seq. No. 3195

Contig ID 2225_1.R1040 5'-most EST zhf700955393.h1

Method BLASTX
NCBI GI g3819099
BLAST score 450
E value 1.0e-44
Match length 172
% identity 49



NCBI Description (AJ009825) copper amine oxidase [Cicer arietinum]

Seq. No. 3196

Contig ID 2226_1.R1040 5'-most EST kl1701205002.h1

Method BLASTX
NCBI GI g2829911
BLAST score 1328
E value 1.0e-147
Match length 363
% identity 70

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]

Seq. No. 3197

Contig ID 2226 2.R1040

5'-most EST LIB3051-067-Q1-K1-E5

Method BLASTX
NCBI GI g2829911
BLAST score 212
E value 9.0e-17
Match length 87
% identity 56

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]

Seq. No. 3198

2226 4.R1040 Contig ID g4291800 5'-most EST Method BLASTX g2829911 NCBI GI BLAST score 788 5.0e-84 E value 206 Match length % identity 44

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]

Seq. No. 3199

Contig ID 2226_5.R1040

5'-most EST jC-gmf102220063d08a1

Method BLASTX
NCBI GI g2829911
BLAST score 296
E value 8.0e-27
Match length 105
% identity 61

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]

Seq. No. 3200

Contig ID 2226 6.R1040

5'-most EST LIB3049-004-Q1-E1-E12

Method BLASTX
NCBI GI g2829911
BLAST score 125
E value 8.0e-12
Match length 77
% identity 57

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]



```
3201
Seq. No.
                   2227 1.R1040 ·
Contig ID
5'-most EST
                   uC-gmropic009b10b1
Method
                   BLASTX
NCBI GI
                   g2065531
BLAST score
                   2076
                   0.0e + 00
E value
Match length
                   484
% identity
                   77
                   (U78526) endo-1,4-beta-glucanase [Lycopersicon esculentum]
NCBI Description
Seq. No.
                   2227 2.R1040
Contig ID
5'-most EST
                   sat701012015.h1
Method
                   BLASTX
NCBI GI
                   g2065531
BLAST score
                   1055
                   1.0e-154
E value
Match length
                   365
% identity
                   72
NCBI Description (U78526) endo-1,4-beta-glucanase [Lycopersicon esculentum]
Seq. No.
                   3203
                   2227 3.R1040
Contig ID
5'-most EST
                   LIB3051-041-Q1-K1-B12
Method
                   BLASTX
NCBI GI
                   g2065531
BLAST score
                   363
                   3.0e - 34
E value
                   92
Match length
% identity
                   (U78526) endo-1,4-beta-glucanase [Lycopersicon esculentum]
NCBI Description
                   3204
Seq. No.
                   2229 1.R1040
Contig ID
                   asn701139633.h1
5'-most EST
                   BLASTN
Method
                   q3366929
NCBI GI
BLAST score
                   35
                   5.0e-10
E value
                   71
Match length
                   87
% identity
                   Solanum tuberosum putative phosphatidylinositol 4-kinase
NCBI Description
                   (PI4KI) mRNA, partial cds
                   3205
Seq. No.
                   2230 1.R1040
Contig ID
                   jex700908180.h1
5'-most EST
Method
                   BLASTX
                   g118158
NCBI GI
BLAST score
                   1697
                   0.0e + 00
E value
Match length
                   362
 % identity
                   87
                   VIGNAIN PRECURSOR (BEAN ENDOPEPTIDASE) (CYSTEINE
NCBI Description
                   PROTEINASE) (SULFHYDRYL-ENDOPEPTIDASE) (SH-EP)
```

>gi_82006_pir__S12581 cysteine proteinase (EC 3.4.22.-) -



black gram >gi_22062_emb_CAA33753_ (X15732) sulfhydryl-pre-endopeptidase (AA -20 to 342) [Vigna mungo] >gi_22066_emb_CAA36181_ (X51900) sulfhydryl-endopeptidase [Vigna mungo]

3206 Seq. No.

2230 2.R1040 Contig ID

5'-most EST jC-gmst02400066a10a2

Seq. No.

2231 1.R1040 Contig ID

uC-qmrominsoy221g11b1 5'-most EST BLASTX Method q2769566 NCBI GI BLAST score 922

2.0e-99 E value 362 Match length 57 % identity

(Y10477) chloroplast thylakoidal processing peptidase NCBI Description

[Arabidopsis thaliana]

3208 Seq. No.

2231 2.R1040 Contig ID rlr700900371.h1 5'-most EST

3209 Seq. No.

2231 4.R1040 Contia ID

LIB3093-038-Q1-K1-G5 5'-most EST

3210 Seq. No.

2231 5.R1040 Contig ID jsh701068480.h1 5'-most EST

BLASTX Method g2769566 NCBI GI BLAST score 246 3.0e-21E value 62 Match length % identity

(Y10477) chloroplast thylakoidal processing peptidase NCBI Description

[Arabidopsis thaliana]

3211 Seq. No.

2234 1.R1040 Contig ID

uC-gmflminsoy071f11b2 5'-most EST

Method BLASTX q729882 NCBI GI 417 BLAST score 1.0e-40 E value Match length 111 % identity

CASEIN KINASE II BETA' CHAIN (CK II) NCBI Description

>gi_1076300_pir__S47968 casein kinase II (EC 2.7.1.-) beta chain CKB2 - Arabidopsis thaliana >gi_467975 (U03984) casein kinase II beta subunit CKB2 [Arabidopsis thaliana] >gi_2245122_emb_CAB10544_ (Z97343) unnamed protein product

[Arabidopsis thaliana]



```
Seq. No.
                  2234 2.R1040
Contig ID
                  zsq701118222.h1
5'-most EST
                  BLASTN
Method
                  g3341671
NCBI GI
                  42
BLAST score
                  3.0e-14
E value
                  122
Match length
% identity
                  84
                  Arabidopsis thaliana chromosome II BAC F16B22 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
                   3213
Seq. No.
                   2237 1.R1040
Contig ID
                   uC-qmrominsoy093g12b1
5'-most EST
                   BLASTX
Method
                   g1171577
NCBI GI
BLAST score
                   745
                   3.0e-79
E value
Match length
                   173
% identity
                   (X95343) hypersensitivity-related gene [Nicotiana tabacum]
NCBI Description
Seq. No.
                   2238 1.R1040
Contiq ID
                   jC-qmf102220114g03a1
5'-most EST
                   3215
Seq. No.
                   2245 1.R1040
Contig ID
                   LIB3028-046-Q1-B1-G11
5'-most EST
                   BLASTX
Method
                   q4455337
NCBI GI
                   290
BLAST score
                   8.0e-26
E value
Match length
                   213
                   31
% identity
                   (AL035525) pectinesterase-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3216
Seq. No.
                   2247 1.R1040
Contig ID
                   LIB3106-021-Q1-K1-C9
5'-most EST
                   BLASTX
Method
NCBI GI
                   g478409
BLAST score
                   491
                   2.0e-73
E value
                   256
Match length
% identity
                   peroxidase (EC 1.11.1.7), cationic - adzuki bean
NCBI Description
```

>gi_218328_dbj_BAA01950_ (D11337) peroxidase [Vigna

angularis]

3217 Seq. No.

2249 1.R1040 Contig ID wrg700789411.h2 5'-most EST

3218 Seq. No.



2250 1.R1040 Contig ID pmv700894288.hl 5'-most EST BLASTX Method g3164115 NCBI GI 2̃31 BLAST score 1.0e-18 E value 133 Match length % identity 37 NCBI Description (AJ224145) major latex-like protein [Rubus idaeus] 3219 Seq. No. 2251 1.R1040 Contig ID LIB3028-046-Q1-B1-G6 5'-most EST 3220 Seq. No. 2253 1.R1040 Contig ID $smc7\overline{0}0748727.h1$ 5'-most EST BLASTX Method g2154609 NCBI GI 390 BLAST score 1.0e-37 E value 136 Match length 51 % identity (D63509) endoxyloglucan transferase related protein NCBI Description [Arabidopsis thaliana] >gi 4522010_gb_AAD21783.1 (AC007069) endoxyloglucan transferase [Arabidopsis thaliana] 3221 Seq. No. 2254 1.R1040 Contig ID LIB3051-045-Q1-K1-C5 5'-most EST BLASTX Method g1019946 NCBI GI BLAST score 852 2.0e-91 E value 210 Match length % identity (U37060) ascorbate peroxidase [Gossypium hirsutum] NCBI Description 3222 Seq. No. 2254 2.R1040 Contig ID 5'-most EST leu701157531.hl Method BLASTX NCBI GI q1019946 BLAST score 531 5.0e-54E value 159 Match length % identity 73 (U37060) ascorbate peroxidase [Gossypium hirsutum] NCBI Description 3223 Seq. No. 2256 1.R1040 Contig ID

LIB3093-027-Q1-K1-A3 5'-most EST

Method BLASTX g4204313 NCBI GI 532 BLAST score 3.0e-54E value



```
Match length
                  85
% identity
                  (AC003027) lcl_prt_seq No definition line found
NCBI Description
                  [Arabidopsis thaliana]
Seq. No.
                  3224
                  2256 2.R1040
Contig ID
                  hyd700727331.h1
5'-most EST
                  BLASTX
Method
                  q4204313
NCBI GI
BLAST score
                  372
                  1.0e-35
E value
                  83
Match length
% identity
                   (AC003027) lcl prt seq No definition line found
NCBI Description
                   [Arabidopsis thaliana]
                  3225
Seq. No.
                  2257 1.R1040
Contig ID
                  ssr700556995.h1
5'-most EST
                  BLASTX
Method
                  q3548808
NCBI GI
                   493
BLAST score
```

2.0e-49 E value 297 Match length 45 % identity

NCBI Description (AC005313) unknown protein [Arabidopsis thaliana]

Seq. No. 3226 Contig ID 2258 1.R1040 jC-gmle01810093f07d1 5'-most EST BLASTN Method NCBI GI q20721 150 BLAST score 1.0e-78 E value

290 Match length 90 % identity

Pea mRNA for ferredoxin-NADP+ reductase NCBI Description

3227 Seq. No.

Contig ID 2258 2.R1040

LIB3049-003-Q1-E1-F4 5'-most EST

Method BLASTX NCBI GI g1351408 BLAST score 1781 0.0e+00E value Match length 467 % identity 79

VACUOLAR PROCESSING ENZYME PRECURSOR (VPE) NCBI Description

>gi 1076563_pir S51117 cystein proteinase - sweet orange >gi_633185_emb_CAA87720_ (Z47793) cystein proteinase (by similarity) [Citrus sinensis] >gi_1588548_prf__2208463A

vascular processing protease [Citrus sinensis]

3228 Seq. No.

2258 3.R1040 Contig ID

uC-gmflminsoy010h10b1 5'-most EST



```
BLASTN
Method
NCBI GI
                   g2511696
                   743
BLAST score
                   0.0e + 00
E value
                   1412
Match length
                   89
% identity
                   Phaseolus vulgaris Moldavian encoding asparagine-specific
NCBI Description
                   endopeptidase precursor (clone cp6b)
                   3229
Seq. No.
                   2258 4.R1040
Contig ID
                   jC-qmle01810009h12d1
5'-most EST
                   BLASTN
Method
```

Method BLASIN
NCBI GI g551130
BLAST score 437
E value 0.0e+00
Match length 877
% identity 88

NCBI Description Vicia faba ferredoxin NADP+ reductase precursor (fnr) mRNA,

complete cds

 Seq. No.
 3230

 Contig ID
 2258_5.R1040

 5'-most EST
 fde700871556.h1

 Method
 BLASTN

 NCBI GI
 g2511696

 BLAST score
 715

NCBI GI g2511696
BLAST score 715
E value 0.0e+00
Match length 1488
% identity 88

NCBI Description Phaseolus vulgaris Moldavian encoding asparagine-specific

endopeptidase precursor (clone cp6b)

Seq. No. 3231

Contig ID 2259 1.R1040 5'-most EST ssr700554396.h1

Method BLASTX
NCBI GI g4539003
BLAST score 203
E value 2.0e-15
Match length 81
% identity 25

NCBI Description (AL049481) putative protein [Arabidopsis thaliana]

Seq. No. 3232

Contig ID 2260_1.R1040 5'-most EST rlr700898755.h1

Method BLASTX
NCBI GI g4544432
BLAST score 227
E value 1.0e-18
Match length 93
% identity 55

NCBI Description (AC006955) putative mannose-1-phosphate guanyltransferase

[Arabidopsis thaliana]

Seq. No. 3233



```
2261 1.R1040
Contig ID
                  LIB3049-053-Q1-E1-E3
5'-most EST
Seq. No.
                  2262 1.R1040
Contig ID
                  uC-gmronoir068e07b1
5'-most EST
Seq. No.
                  3235
                  2262 2.R1040
Contig ID
                  LIB3170-078-Q1-K1-A7
5'-most EST
                  BLASTX
Method
                  g4049344
NCBI GI
                  158
BLAST score
                  3.0e-10
E value
                  214
Match length
                  11
% identity
NCBI Description (AL034567) putative protein [Arabidopsis thaliana]
                  3236
Seq. No.
                  2262 3.R1040
Contig ID
                  uC-gmflminsoy089e10b1
5'-most EST
                   3237
Seq. No.
                   2264 1.R1040
Contig ID
                  LIB3138-074-P1-N1-D5
5'-most EST
                   3238
Seq. No.
                   2268 1.R1040
Contig ID
                   jC-gmst02400072c10a1
5'-most EST
                   BLASTX
Method
                   g3776005
NCBI GI
BLAST score
                   663
                   1.0e-69
E value
Match length
                   156
                   87
% identity
                   (AJ010466) RNA helicase [Arabidopsis thaliana]
NCBI Description
                   3239
Seq. No.
Contig ID
                   2268 2.R1040
                   LIB3050-002-Q1-E1-H10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3776005
BLAST score
                   444
E value
                   3.0e-44
                   94
Match length
                   93
% identity
                   (AJ010466) RNA helicase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3240
                   2269 1.R1040
Contig ID
                   LIB3109-011-Q1-K1-E11
5'-most EST
Method
                   BLASTN
NCBI GI
                   g1666229
BLAST score
                   327
```

0.0e + 00

655

87

E value

Match length

% identity



NCBI Description Pisum sativum actin (PEAc9) mRNA, complete cds

>gi_1724138_gb_U81047_PSU81047 Pisum sativum actin (PEAc9)

mRNA, complete cds

Seq. No. 3241

Contig ID 2269_2.R1040

5'-most EST jC-gmst02400043b09a1

Method BLASTN
NCBI GI g1666229
BLAST score 358
E value 0.0e+00
Match length 719
% identity 89

NCBI Description Pisum sativum actin (PEAc9) mRNA, complete cds

>gi_1724138_gb_U81047_PSU81047 Pisum sativum actin (PEAc9)

mRNA, complete cds

Seq. No. 3242

Contig ID 2269_3.R1040 5'-most EST ncj700986316.h1

Method BLASTN
NCBI GI g940812
BLAST score 167
E value 1.0e-88
Match length 394
% identity 86

NCBI Description P.sativum mRNA for actin protein

Seq. No. 3243

Contig ID 2269_4.R1040 5'-most EST wvk700683029.h1

Method BLASTN
NCBI GI g3860316
BLAST score 192
E value 1.0e-103
Match length 465
% identity 88

NCBI Description Cicer arietinum mRNA for actin, partial

Seq. No. 3244

Contig ID 2269_5.R1040 5'-most EST zhf700963064.h1

Method BLASTN
NCBI GI g18317
BLAST score 113
E value 7.0e-57
Match length 197
% identity 89

NCBI Description D.carota RNA for actin gene Ac1

Seq. No. 3245

Contig ID 2270_1.R1040 5'-most EST rca700997077.h1

Seq. No. 3246

Contig ID 2270_2.R1040

5'-most EST LIB3074-036-Q1-K1-F11



BLASTX Method g1881585 NCBI GI 330 BLAST score 2.0e-30 E value 145 Match length 52 % identity NCBI Description (U72489) remorin [Solanum tuberosum] 3247 Seq. No. 2271 1.R1040 Contig ID fC-gmro700848271a1 5'-most EST BLASTX Method g1905785 NCBI GI 895 BLAST score 2.0e-96 E value

60 % identity NCBI Description (Y10685) G/HBF-1 [Glycine max]

3248 Seq. No.

Match length

2271 2.R1040 Contig ID

uC-gmflminsoy020d04b1 5'-most EST

347

BLASTX Method g1905785 NCBI GI 154 BLAST score 2.0e-15 E value 69 Match length 75 % identity

NCBI Description (Y10685) G/HBF-1 [Glycine max]

Seq. No. 3249

2272 1.R1040 Contig ID 5'-most EST fde700870883.h1

3250 Seq. No.

2277 1.R1040 Contig ID

LIB3093-058-Q1-K1-E12 5'-most EST

Method BLASTX q3075488 NCBI GI BLAST score 1240 1.0e-137 E value Match length 286 % identity

(AF058796) chlorophyll a/b-binding protein [Oryza sativa] NCBI Description

Seq. No.

2277 2.R1040 Contig ID

5'-most EST jC-gmst02400014b12d1

Method BLASTN g3075487 NCBI GI BLAST score 74 5.0e-33 E value Match length 180 % identity 79

Oryza sativa chlorophyll a/b-binding protein (RCABP69) NCBI Description

mRNA, complete cds



```
3252
Seq. No.
Contig ID
                  2277 3.R1040
5'-most EST
                  LIB3138-062-Q1-N1-G10
Method
                  BLASTX
                  g2827649
NCBI GI
BLAST score
                  156
                  3.0e-10
E value
Match length
                  46
% identity
                  61
                  (AL021637) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  3253
Contig ID
                  2277 4.R1040
5'-most EST
                  LIB3065-008-Q1-N1-A8
Method
                  BLASTX
NCBI GI
                  g3075488
BLAST score
                  486
E value
                  8.0e-49
Match length
                  119
% identity
                  84
NCBI Description (AF058796) chlorophyll a/b-binding protein [Oryza sativa]
Seq. No.
                  3254
                  2277 5.R1040
Contig ID
5'-most EST
                  LIB3093-019-Q1-K2-H3
Method
                  BLASTN
NCBI GI
                  g3075487
BLAST score
                  58
                  7.0e-24
E value
Match length
                  86
% identity
                  92
NCBI Description
                  Oryza sativa chlorophyll a/b-binding protein (RCABP69)
                  mRNA, complete cds
                  3255
Seq. No.
Contig ID
                  2278 1.R1040
5'-most EST
                  LIB3107-053-Q1-K1-E8
Method
                  BLASTX
NCBI GI
                  g2244970
BLAST score
                  2257
E value
                  0.0e + 00
Match length
                  585
                   69
% identity
NCBI Description
                  (Z97340) hypothetical protein [Arabidopsis thaliana]
                  >gi 2326365 emb CAA74765 (Y14423) putative cell wall
                  protein [Arabidopsis thaliana]
Seq. No.
                   3256
Contig ID
                  2280 1.R1040
5'-most EST
                  ssr700556839.h1
Method
                  BLASTX
NCBI GI
                  q1710807
BLAST score
                  2464
```

0.0e+00

584 86

E value Match length

% identity



KD CHAPERONIN ALPHA SUBUNIT) (CPN-60 ALPHA) >gi_1185390 (U21105) alphacpn60 [Pisum sativum]

3257 Seq. No. 2280 2.R1040 Contig ID LIB3093-027-Q1-K1-D2 5'-most EST BLASTN Method NCBI GI g3790440 256 BLAST score 1.0e-142 E value 367 Match length 92 % identity Canavalia lineata chaperonin 60 alpha subunit mRNA, nuclear NCBI Description gene encoding chloroplast protein, complete cds Seq. No. 2280 4.R1040 Contig ID epx701109229.h1 5'-most EST Method BLASTN g3790440 NCBI GI BLAST score 84 2.0e-39 E value 272 Match length % identity 83 Canavalia lineata chaperonin 60 alpha subunit mRNA, nuclear NCBI Description gene encoding chloroplast protein, complete cds 3259 Seq. No. 2280 5.R1040 Contig ID LIB3051-052-Q1-K1-A10 5'-most EST BLASTN Method g3790440 NCBI GI 52 BLAST score 2.0e-20 E value 152 Match length 84 % identity Canavalia lineata chaperonin 60 alpha subunit mRNA, nuclear NCBI Description gene encoding chloroplast protein, complete cds 3260 Seq. No. 2281 1.R1040 Contig ID wrg700791180.h1 5'-most EST BLASTX Method NCBI GI g2136011 BLAST score 228 2.0e-18 E value 84 Match length % identity

NCBI Description protein DS 1, 24K - human >gi_1045059_emb_CAA57387_

(X81788) ICT1 protein [Homo sapiens]

>gi_4557657_ref_NP_001536.1_pICT1_ immature colon carcinoma

transcript

Seq. No. 3261

Contig ID 2284_1.R1040

5'-most EST fC-gmro700835781b1

Method BLASTX



NCBI GI g3482919 BLAST score 900 E value 4.0e-97 Match length 241 % identity 71

NCBI Description (AC003970) Putative protein kinase [Arabidopsis thaliana]

Seq. No. 3262

2290 1.R1040 Contig ID q5605900 5'-most EST BLASTX Method q2253442 NCBI GI 270 BLAST score 3.0e-23 E value 78 Match length 56 % identity

NCBI Description (AF007784) LTCOR11 [Lavatera thuringiaca]

Seq. No. 3263

Contig ID 2290_2.R1040

5'-most EST LIB3028-023-Q1-B1-A9

Seq. No. 3264

Contig ID 2290_3.R1040 5'-most EST jex700904808.h1

Seq. No. 3265

Contig ID 2290_4.R1040 5'-most EST crh700851666.h1

Method BLASTX
NCBI GI g2253442
BLAST score 191
E value 1.0e-14
Match length 46
% identity 63

NCBI Description (AF007784) LTCOR11 [Lavatera thuringiaca]

Seq. No. 3266

Contig ID 2291_1.R1040

5'-most EST LIB3049-008-Q1-E1-H2

Method BLASTX
NCBI GI g3298548
BLAST score 340
E value 6.0e-32
Match length 138
% identity 63

NCBI Description (AC004681) putative spliceosomal protein [Arabidopsis

thaliana]

Seq. No. 3267

Contig ID 2292_1.R1040

5'-most EST LIB3051-011-Q1-E1-G12

Seq. No. 3268

Contig ID 2292 2.R1040

5'-most EST LIB3170-085-Q1-K1-D6



```
3269
Seq. No.
                  2293 1.R1040
Contig ID
                  q5175583
5'-most EST
                  BLASTX
Method
                  q2078350
NCBI GI
                  799
BLAST score
                   3.0e-85
E value
                  192
Match length
                   83
% identity
NCBI Description (U95923) transaldolase [Solanum tuberosum]
                   3270
Seq. No.
                   2293 2.R1040
Contig ID
                  LIB3139-031-P1-N1-G9
5'-most EST
Method
                   BLASTX
                   g2078350
NCBI GI
                   1457
BLAST score
                   1.0e-162
E value
                   384
Match length
                   75
% identity
NCBI Description (U95923) transaldolase [Solanum tuberosum]
                   3271
Seq. No.
                   2293 4.R1040
Contig ID
                   uC-qmrominsoy026g11b1
5'-most EST
                   BLASTX
Method
                   q2078350
NCBI GI
                   321
BLAST score
                   1.0e-29
E value
Match length
                   71
% identity
                   83
                   (U95923) transaldolase [Solanum tuberosum]
NCBI Description
                   3272
Seq. No.
                   2294 1.R1040
Contig ID
                   LIB3028-009-Q1-B1-G4
5'-most EST
                   BLASTX
Method
                   g1935019
NCBI GI
                   438
BLAST score
                   4.0e-43
E value
                   161
Match length
                   56
 % identity
                   (Z93774) sucrose transport protein [Vicia faba]
NCBI Description
                   3273
 Seq. No.
                   2294 2.R1040
 Contig ID
                   smc700749544.hl
 5'-most EST
                   BLASTX
 Method
                   g1935019
 NCBI GI
 BLAST score
                   473
                   2.0e-47
 E value
                   112
 Match length
 % identity
                   (Z93774) sucrose transport protein [Vicia faba]
 NCBI Description
```

668

3274

2294_3.R1040

Seq. No.

Contig ID



```
LIB3093-015-Q1-K1-G9
5'-most EST
                  3275
Seq. No.
                  2296 1.R1040
Contig ID
                  kl1701214421.hl
5'-most EST
                  BLASTX
Method
                  g2245135
NCBI GI
BLAST score
                  581
                  7.0e-60
E value
Match length
                  156
                  70
% identity
NCBI Description (Z97344) hypothetical protein [Arabidopsis thaliana]
                  3276
Seq. No.
                  2297 1.R1040
Contig ID
                  LIB3107-025-Q1-K1-G8
5'-most EST
                  BLASTX
Method
                   g4433048
NCBI GI
BLAST score
                  592
                   6.0e-61
E value
                  147
Match length
                   79
% identity
NCBI Description (D26578) DNA-binding protein [Daucus carota]
Seq. No.
                   3277
                   2297 2.R1040
Contig ID
                   uC-gmflminsoy064a03b1
5'-most EST
                   BLASTX
Method
                   g4433048
NCBI GI
                   592
BLAST score
                   7.0e-61
E value
                   154
Match length
                   78
% identity
NCBI Description (D26578) DNA-binding protein [Daucus carota]
                   3278
Seq. No.
                   2297 5.R1040
Contig ID
5'-most EST
                   zsg701125668.hl
                   3279
Seq. No.
                   2298 1.R1040
Contig ID
                   uC-gmropic094g04b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2827143
                   937
BLAST score
                   1.0e-101
E value
                   195
Match length
                   89
 % identity
                   (AF027174) cellulose synthase catalytic subunit
NCBI Description
                   [Arabidopsis thaliana]
                   3280
 Seq. No.
                   2298 2.R1040
 Contig ID
 5'-most EST
                   g4292031
 Method
                   BLASTX
```

g2827143

415

NCBI GI

BLAST score



```
1.0e-40
E value
Match length
                  79
                  92
% identity
                   (AF027174) cellulose synthase catalytic subunit
NCBI Description
                   [Arabidopsis thaliana]
                   3281
Seq. No.
                   2298 3.R1040
Contig ID
5'-most EST
                   dpv701097183.hl
                  BLASTX
Method
                   a2827143
NCBI GI
                   2504
BLAST score
                   0.0e + 00
E value
                   522
Match length
                   89
% identity
                   (AF027174) cellulose synthase catalytic subunit
NCBI Description
                   [Arabidopsis thaliana]
                   3282
Seq. No.
                   2305 1.R1040
Contig ID
                   pxt700946213.h1
5'-most EST
                   BLASTX
Method
                   q4185513
NCBI GI
                   589
BLAST score
                   1.0e-60
E value
                   130
Match length
% identity
                   (AF102823) actin depolymerizing factor 5 [Arabidopsis
NCBI Description
                   thaliana] >gi_4185517 (AF102825) actin depolymerizing
                   factor 5 [Arabidopsis thaliana]
                   3283
Seq. No.
                   2305 2.R1040
Contig ID
                   LIB3107-025-Q1-K1-C12
5'-most EST
Method
                   BLASTX
                   g4185513
NCBI GI
                   613
BLAST score
                   1.0e-63
E value
                   143
Match length
                   83
 % identity
                   (AF102823) actin depolymerizing factor 5 [Arabidopsis
NCBI Description
                   thaliana] >gi_4185517 (AF102825) actin depolymerizing
                   factor 5 [Arabidopsis thaliana]
 Seq. No.
                   3284
                   2306 1.R1040
 Contig ID
                   LIB3051-042-Q1-K1-H9
 5'-most EST
```

Method BLASTX
NCBI GI g2262195
BLAST score 323
E value 1.0e-29
Match length 150
% identity 42

NCBI Description (U64820) josephin MJD1 [Homo sapiens]

Seq. No. 3285

Contig ID 2306_3.R1040



5'-most EST LIB3051-104-Q1-K1-G1

Method BLASTX
NCBI GI g1781299
BLAST score 397
E value 2.0e-38
Match length 151
% identity 57

NCBI Description (Y09506) transformer-SR ribonucleoprotein [Nicotiana

tabacum]

Seq. No. 3286

2306 4.R1040 Contig ID q5687810 5'-most EST Method BLASTX NCBI GI q1781299 BLAST score 272 8.0e-28 E value Match length 110 % identity 66

NCBI Description (Y09506) transformer-SR ribonucleoprotein [Nicotiana

tabacum]

Seq. No. 3287

Contig ID 2307_1.R1040 5'-most EST ncj700980603.h1

Seq. No. 3288

Contig ID 2309 2.R1040

5'-most EST uC-gmrominsoy074e05b1

Seq. No. 3289

Contig ID 2310 1.R1040 5'-most EST jex700908408.h1

Seq. No. 3290

Contig ID 2310 2.R1040 5'-most EST kl1701202645.h1

Seq. No. 3291

Contig ID 2311 1.R1040

5'-most EST uC-gmflminsoy061a11b1

Method BLASTX
NCBI GI g2511693
BLAST score 2060
E value 0.0e+00
Match length 458
% identity 83

NCBI Description (Z99954) cysteine proteinase precursor [Phaseolus vulgaris]

Seq. No. 3292

Contig ID 2311_2.R1040 5'-most EST taw700658011.h1

Method BLASTN
NCBI GI g2511692
BLAST score 168
E value 3.0e-89
Match length 454



% identity NCBI Description

Phaseolus vulgaris Moldavian encoding cysteine proteinase precursor (clone cp71)

Seq. No.

3293

Contig ID 5'-most EST 2311 3.R1040 zhf700958286.h1

Method NCBI GI BLASTN g2511692

BLAST score

70 4.0e-31

E value Match length

209

% identity

86

NCBI Description

Phaseolus vulgaris Moldavian encoding cysteine proteinase

precursor (clone cp71)

Seq. No.

3294

Contig ID 5'-most EST

2311 5.R1040 xpa700793534.h1

Method NCBI GI BLASTN g2511692

BLAST score

55

E value

4.0e-22

Match length % identity

170 86

NCBI Description

Phaseolus vulgaris Moldavian encoding cysteine proteinase

precursor (clone cp71)

Seq. No.

3295

Contig ID 5'-most EST 2312 1.R1040 dpv701103305.h1

Method NCBI GI BLASTX g1350720

BLAST score E value

452 1.0e-44

Match length % identity

111 76

NCBI Description

60S RIBOSOMAL PROTEIN L32

Seq. No.

3296

Contig ID

2312 2.R1040

5'-most EST

LIB3028-045-Q1-B1-H12

Method NCBI GI BLASTX g1350720

BLAST score E value

452 6.0e-45

Match length

111

% identity NCBI Description

60S RIBOSOMAL PROTEIN L32

Seq. No.

3297

76

Contig ID

2312 3.R1040

5'-most EST

LIB3040-060-Q1-E1-E11

Method BLASTX NCBI GI q1350720 BLAST score 438 E value 3.0e-43



Match length 94 % identity 85

NCBI Description 60S RIBOSOMAL PROTEIN L32

Seq. No. 3298

Contig ID 2312_4.R1040

5'-most EST LIB3073-013-Q1-K1-C6

Method BLASTX
NCBI GI g132886
BLAST score 246
E value 1.0e-20
Match length 130
% identity 43

NCBI Description 60S RIBOSOMAL PROTEIN L32 >gi_71335_pir_R5HU32 ribosomal

protein L32 - human >gi_71336_pir__R5MS32 ribosomal protein L32 - mouse >gi_71337_pir__R5RT32 ribosomal protein L32 - rat >gi_36132_emb_CAA27048 (X03342) rpL32 (aa 1-135) [Homo sapiens] >gi_57117_emb_CAA29777_ (X06483) ribosomal protein

L32 [Rattus norvegicus] >gi_200781 (K02060) ribosomal protein L32-3A [Mus musculus] >gi_226004_prf__1405339A

ribosomal protein L32 [Rattus norvegicus]

>gi 4506635 ref_NP_000985.1_pRPL32_ ribosomal protein L32

Seq. No. 3299

Contig ID 2312 5.R1040

5'-most EST LIB3106-086-Q1-K1-H6

Method BLASTX
NCBI GI g132886
BLAST score 185
E value 6.0e-14
Match length 54
% identity 67

NCBI Description

60S RIBOSOMAL PROTEIN L32 >gi_71335_pir_ R5HU32 ribosomal protein L32 - human >gi_71336_pir_ R5MS32 ribosomal protein L32 - mouse >gi_71337_pir_ R5RT32 ribosomal protein L32 - rat >gi_36132_emb_CAA27048 (X03342) rpL32 (aa 1-135) [Homo sapiens] >gi_57117_emb_CAA29777 (X06483) ribosomal protein L32 [Rattus norvegicus] >gi_200781 (K02060) ribosomal protein L32-3A [Mus musculus] >gi_226004_prf__1405339A

ribosomal protein L32 [Rattus norvegicus]

>gi_4506635_ref_NP_000985.1_pRPL32_ ribosomal protein L32

Seq. No. 3300

Contig ID 2313_1.R1040

5'-most EST LIB3049-045-Q1-E1-G8

Method BLASTX
NCBI GI g3608485
BLAST score 1089
E value 1.0e-119
Match length 223
% identity 92

NCBI Description (AF088915) proteasome beta subunit [Petunia x hybrida]

Seq. No. 3303

Contig ID 2313_2.R1040 5'-most EST zzp700835212.h1

Method BLASTX



```
g3608485
NCBI GI
BLAST score
                   455
E value
                   2.0e-45
Match length
                   92
% identity
                   90
NCBI Description
                  (AF088915) proteasome beta subunit [Petunia x hybrida]
                   3302
Seq. No.
Contig ID
                  2315 1.R1040
5'-most EST
                   g4292157
Method
                  BLASTX
NCBI GI
                   g2347088
BLAST score
                   417
E value
                   2.0e-40
Match length
                  117
% identity
                   70
                   (U72765) non-specific lipid transfer protein PvLTP-24
NCBI Description
                   [Phaseolus vulgaris]
Seq. No.
                  3303
                  2315 2.R1040
Contig ID
5'-most EST
                  g5753304
Method
                  BLASTN
NCBI GI
                  q4490324
BLAST score
                   40
E value
                   5.0e-13
Match length
                   88
% identity
                   86
                  Arabidopsis thaliana DNA chromosome 4, BAC clone T9A14
NCBI Description
                   (ESSA project)
Seq. No.
                   3304
Contig ID
                   2315 3.R1040
5'-most EST
                  g5126391
Method
                  BLASTX
NCBI GI
                   g4490331
BLAST score
                   520
E value
                  1.0e-52
Match length
                  139
% identity
                  70
NCBI Description
                  (AL035656) hypothetical protein [Arabidopsis thaliana]
                   3305
Seq. No.
Contig ID
                   2316 1.R1040
5'-most EST
                   leu701157155.h1
Method
                  BLASTX
NCBI GI
                   g2160182
BLAST score
                   206
E value
                   9.0e-16
                  155
Match length
                   37
% identity
```

NCBI Description

(AC000132) ESTs gb ATTS1236, gb T43334, gb N97019, gb AA395203

come from this gene. [Arabidopsis thaliana]

3306 Seq. No.

Contig ID 2316 2.R1040

5'-most EST LIB3170-012-Q1-J1-A10



Seq. No. 3307

Contig ID 2316 3.R1040

5'-most EST LIB3039-012-Q1-E1-C9

Seq. No. 3308

Contig ID 2317_1.R1040

5'-most EST uC-gmropic007e02b1

Method BLASTX
NCBI GI g4309738
BLAST score 820
E value 1.0e-87
Match length 297
% identity 58

NCBI Description (AC006439) putative tubby protein [Arabidopsis thaliana]

Seq. No. 3309

Contig ID 2317_2.R1040 5'-most EST vwf700673985.h1

Method BLASTX
NCBI GI g4309738
BLAST score 505
E value 5.0e-51
Match length 145
% identity 66

NCBI Description (AC006439) putative tubby protein [Arabidopsis thaliana]

Seq. No. 3310

Contig ID 2321 1.R1040

5'-most EST LIB3028-046-Q1-B1-A11

Method BLASTX
NCBI GI g1263160
BLAST score 360
E value 6.0e-34
Match length 251
% identity 12

NCBI Description (X89226) leucine-rich repeat/receptor protein kinase [Oryza

sativa]

Seq. No. 3311

Contig ID 2322_1.R1040 5'-most EST asn701130506.h1

Method BLASTN
NCBI GI g1370195
BLAST score 308
E value 1.0e-172
Match length 412
% identity 94

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB8D

Seq. No. 3312

Contig ID 2322 2.R1040 5'-most EST sat701004249.h1

Method BLASTN
NCBI GI g1370195
BLAST score 256
E value 1.0e-142



Match length 497 % identity 89

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB8D

Seq. No. 3313

Contig ID 2322_3.R1040

5'-most EST LIB3087-002-Q1-K1-F6

Method BLASTN
NCBI GI g1370197
BLAST score 294
E value 1.0e-164
Match length 402
% identity 93

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB8E

Seq. No. 3314

Contig ID 2322_4.R1040

5'-most EST fC-gmse700751577a1

Method BLASTN
NCBI GI g1370195
BLAST score 210
E value 1.0e-114
Match length 302
% identity 92

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB8D

Seq. No. 3315

Contig ID 2323_1.R1040

5'-most EST LIB3092-061-Q1-K1-G5

Method BLASTX
NCBI GI g2495091
BLAST score 362
E value 7.0e-34
Match length 175
% identity 42

NCBI Description GRPE PROTEIN >gi_1075599_pir__PC2235 GrpE protein -

Synechococcus sp. (PCC 7942) (fragment)

>gi_507817_dbj_BAA05902_ (D28550) heat shock protein GrpE

homolog [Synechococcus sp.]

Seq. No. 3316

Contig ID 2326_1.R1040

5'-most EST jC-gmst02400073c02a1

Method BLASTX
NCBI GI g3005599
BLAST score 223
E value 4.0e-18
Match length 168
% identity 28

NCBI Description (AF052432) katanin p80 subunit [Homo sapiens]

Seq. No. 3317

Contig ID 2326_2.R1040 5'-most EST crh700853505.h1

Method BLASTX
NCBI GI g940288
BLAST score 778

```
E value
                   1.0e-82
Match length
                   223
% identity
                   68
                   (L43510) protein localized in the nucleoli of pea nuclei;
NCBI Description
                  ORF; putative [Pisum sativum]
                   3318
Seq. No.
Contig ID
                  2328 1.R1040
                  g5677260
5'-most EST
Method
                  BLASTX
                  g2244910
NCBI GI
BLAST score
                  316
```

E value 6.0e-29 Match length 105 50 % identity

NCBI Description (Z97339) unnamed protein product [Arabidopsis thaliana]

Seq. No. 3319 2329 1.R1040 Contig ID 5'-most EST

LIB3028-045-Q1-B1-F2

BLASTX Method NCBI GI q2832695 BLAST score 181 E value 3.0e-13 Match length 62 % identity 58

(AL021713) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 3320

Contig ID 2330 1.R1040 5'-most EST $k117\overline{0}1202518.h1$

Method BLASTX NCBI GI g1518855 BLAST score 211 2.0e-16 E value Match length 72 % identity 61

NCBI Description (U65960) HASPP28 [Homo sapiens]

3321 Seq. No.

2330 2.R1040 Contig ID

5'-most EST LIB3109-027-Q1-K1-E1

Method BLASTN NCBI GI g3821780 BLAST score 36 1.0e-10 E value Match length 48 % identity 65

NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No.

Contig ID 2330 4.R1040

5'-most EST LIB3051-024-Q1-K1-H6

3322

Seq. No. 3323

2330 5.R1040 Contig ID

5'-most EST LIB3040-015-Q1-E1-A2



```
3324
Seq. No.
                   2330 7.R1040
Contig ID
                   gsv701055923.hl
5'-most EST
                   3325
Seq. No.
                   2333 1.R1040
Contig ID
                   ssr700559590.hl
5'-most EST
                   BLASTX
Method
                   q4469011
NCBI GI
                   881
BLAST score
                   6.0e-95
E value
                   238
Match length
                   74
% identity
                   (AL035602) carbohydrate kinase-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3326
Seq. No.
                   2333 2.R1040
Contig ID
                   LIB3106-097-Q1-K1-F11
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4469011
                   223
BLAST score
                   5.0e-18
E value
Match length
                   68
                   87
% identity
                   (AL035602) carbohydrate kinase-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3327
Seq. No.
                   2334 1.R1040
Contig ID
                   ncj700985896.h1
5'-most EST
                   BLASTX
Method
                   g4455284
NCBI GI
                   617
BLAST score
                   4.0e-64
E value
                   163
Match length
                   69
% identity
                   (AL035527) beta-glucosidase-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3328
Seq. No.
                   2334 2.R1040
Contig ID
5'-most EST
                   LIB3028-041-Q1-B1-F2
Method
                   BLASTX
                   q4455284
NCBI GI
                   395
BLAST score
                   2.0e-38
E value
Match length
                   109
% identity
                    68
                    (AL035527) beta-glucosidase-like protein [Arabidopsis
NCBI Description
                   thaliana]
```

Seq. No. 3329

Contig ID 2339_1.R1040

5'-most EST LIB3055-005-Q1-N1-C11

Method BLASTX

BLAST score

Match length

% identity

E value

85

216 87

7.0e-40



```
q1710840
NCBI GI
                  2261
BLAST score
                  0.0e+00
E value
                  480
Match length
                  89
% identity
                  ADENOSYLHOMOCYSTEINASE (S-ADENOSYL-L-HOMOCYSTEINE
NCBI Description
                  HYDROLASE) (ADOHCYASE) (CYTOKININ BINDING PROTEIN CBP57)
                  >gi 441217_dbj_BAA03709_ (D16138) S-adenosyl-L-homocystein
                  hydrolase [Nicotiana sylvestris] >gi_1857024_dbj_BAA08142_
                  (D45204) S-adenosyl-L-homocysteine hydrolase [Nicotiana
                  tabacum] >gi_2588781_dbj_BAA23164_ (D49804)
                  S-adenosyl-L-homocysteine hydrolase [Nicotiana tabacum]
                  3330
Seq. No.
                  2339 3.R1040
Contig ID
                  zhf700951722.h1
5'-most EST
                  BLASTN
Method
                  q170772
NCBI GI
BLAST score
                  176
                   4.0e-94
E value
                  296
Match length
                   90
% identity
                  Triticum aestivum S-adenosyl-L-homocysteine hydrolase
NCBI Description
                   (SH6.2) mRNA, complete cds
                   3331
Seq. No.
                   2339 4.R1040
Contig ID
                   sat701008057.h1
5'-most EST
                   BLASTN
Method
                   g535583
NCBI GI
                   110
BLAST score
                   6.0e-55
E value
                   206
Match length
                   88
% identity
                   Medicago sativa adenosylhomocysteinase mRNA, complete cds
NCBI Description
                   3332
Seq. No.
                   2339 5.R1040
Contig ID
                   kl1701209855.h1
 5'-most EST
                   BLASTN
Method
                   g2244747
NCBI GI
                   60
BLAST score
                   4.0e-25
E value
                   124
Match length
 % identity
                   88
                   Arabidopsis thaliana DNA chromosome 4, ESSA I contig
 NCBI Description
                   fragment No
                   3333
 Seq. No.
                   2339 6.R1040
 Contig ID
 5'-most EST
                   LIB3139-004-P1-N1-H3
                   BLASTN
 Method
                   q535583
 NCBI GI
```



NCBI Description Medicago sativa adenosylhomocysteinase mRNA, complete cds

Seq. No. 3334

Contig ID 2339_9.R1040 5'-most EST txt700735356.h1

Method BLASTN
NCBI GI g170772
BLAST score 74
E value 1.0e-33
Match length 102
% identity 93

NCBI Description Triticum aestivum S-adenosyl-L-homocysteine hydrolase

(SH6.2) mRNA, complete cds

Seq. No. 3335

Contig ID 2341_1.R1040 5'-most EST seb700649714.h1

Seq. No. 3336

Contig ID 2341 2.R1040 5'-most EST txt700734826.h1

Seq. No. 3337

Contig ID 2342_1.R1040 5'-most EST kll701213196.h1

Method BLASTX
NCBI GI g2780365
BLAST score 229
E value 8.0e-19
Match length 90
% identity 52

NCBI Description (AB007693) Elongin C [Drosophila melanogaster]

Seq. No. 3338

Contig ID 2343 1.R1040

5'-most EST LIB3\overline{1}70-082-Q1-K1-G9

Method BLASTX
NCBI GI g1172874
BLAST score 721
E value 5.0e-76
Match length 291
% identity 57

NCBI Description DEHYDRATION-RESPONSIVE PROTEIN RD22 PRECURSOR

>gi_479589_pir__S34823 dehydration-induced protein RD22 Arabidopsis thaliana >gi_391608_dbj_BAA01546_ (D10703) rd22
[Arabidopsis thaliana] >gi_447134_prf__1913421A rd22 gene

[Arabidopsis thaliana]

Seq. No. 3339

Contig ID 2343_3.R1040

5'-most EST LIB3170-082-Q1-J1-G9

Method BLASTX
NCBI GI g1172874
BLAST score 300
E value 3.0e-27
Match length 70
% identity 74



DEHYDRATION-RESPONSIVE PROTEIN RD22 PRECURSOR NCBI Description >gi 479589 pir S34823 dehydration-induced protein RD22 -Arabidopsis thaliana >gi_391608_dbj_BAA01546_ (D10703) rd22 [Arabidopsis thaliana] $> gi_4471\overline{3}4_prf_19134\overline{2}1A$ rd22 gene [Arabidopsis thaliana] Seq. No. 3340 2344 1.R1040 Contig ID fua701038156.h1 5'-most EST BLASTX Method g3894158 NCBI GI BLAST score 420 7.0e-41 E value 198 Match length 43 % identity (AC005312) similar to phloem-specific lectin [Arabidopsis NCBI Description thaliana] 3341 Seq. No. 2346 1.R1040 Contig ID jC-gmf102220054g02a1 5'-most EST 3342 Seq. No. 2348 1.R1040 Contig ID LIB3167-010-P1-K1-F5 5'-most EST Method BLASTX g1321941 NCBI GI BLAST score 938 1.0e-101 E value 330 Match length 59 % identity (Z48564) dihydrolipoamide dehydrogenase [Synechocystis NCBI Description PCC6803] 3343 Seq. No. 2349 1.R1040 Contig ID LIB3028-037-Q1-B1-G11 5'-most EST 3344 Seq. No. 2352 1.R1040 Contig ID 5'-most EST uC-qmflminsoy043b08b1 Method BLASTN NCBI GI g871505

BLAST score 452 0.0e + 00E value 612 Match length % identity 94

P.sativum mRNA for small GTP-binding protein (clone pGTP11) NCBI Description

3345 Seq. No.

2352 2.R1040 Contig ID

LIB3049-012-Q1-E1-B1 5'-most EST

Method BLASTN q871507 NCBI GI 474 BLAST score 0.0e+00E value 808 Match length



```
% identity
NCBI Description P.sativum mRNA for small GTP-binding protein (clone pGTP13)
                   3346
Seq. No.
                   2352 4.R1040
Contig ID
5'-most EST
                   asn7\overline{0}1139218.h1
Method
                   BLASTN
                   g871507
NCBI GI
                   218
BLAST score
                   1.0e-119
E value
                   302
Match length
                   93
% identity
                   P.sativum mRNA for small GTP-binding protein (clone pGTP13)
NCBI Description
                   3347
Seq. No.
                   2353 1.R1040
Contig ID
5'-most EST
                   jC-gmst02400072a07a1
Method
                   BLASTX
                   g4415908
NCBI GI
BLAST score
                   542
                   4.0e-69
E value
                   271
Match length
% identity
                   60
                   (AC006282) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3348
Seq. No.
                   2354 1.R1040
Contig ID
                   kl1701214239.h1
5'-most EST
                   3349
Seq. No.
                   2356 1.R1040
Contig ID
5'-most EST
                   LIB3028-045-Q1-B1-E6
                   3350
Seq. No.
Contig ID
                   2357 1.R1040
5'-most EST
                   kl1701206565.h1
Seq. No.
                   3351
                   2357 2.R1040
Contig ID
                   LIB3\overline{1}38-122-Q1-N1-D7
5'-most EST
Seq. No.
                   3352
Contig ID
                   2357 3.R1040
                   LIB3028-045-Q1-B1-E7
5'-most EST
Seq. No.
                   3353
```

2358 1.R1040 Contig ID LIB3050-018-Q1-E1-B10 5'-most EST

BLASTX Method g4510363 NCBI GI 556 BLAST score 7.0e-57 E value Match length 143 73 % identity

(AC007017) putative DNA-binding protein [Arabidopsis NCBI Description

thaliana]

% identity

Seq. No.

32

3359

NCBI Description (U64906) ATFP3 [Arabidopsis thaliana]



```
Seq. No.
                  2358 2.R1040
Contig ID
                  LIB3028-045-Q1-B1-E8
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4510363
BLAST score
                  320
                  1.0e-29
E value
Match length
                  72
                  82
% identity
                  (AC007017) putative DNA-binding protein [Arabidopsis
NCBI Description
                  thaliana]
                  3355
Seq. No.
Contig ID
                  2364 1.R1040
                  LIB3028-010-Q1-B1-E4
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3687240
BLAST score
                  154
                   4.0e-10
E value
Match length
                  40
% identity
                  (AC005169) extensin-like protein [Arabidopsis thaliana]
NCBI Description
                   3356
Seq. No.
                   2365 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220082f04d1
Method
                   BLASTX
                   q4510383
NCBI GI
BLAST score
                   569
E value
                   3.0e-73
Match length
                   253
                   63
% identity
                  (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3357
                   2368 1.R1040
Contig ID
5'-most EST
                   LIB3028-045-Q1-B1-C7
Method
                   BLASTX
                   g3063458
NCBI GI
BLAST score
                   215
E value
                   2.0e-17
Match length
                   80
% identity
                   51
                  (AC003981) F22013.20 [Arabidopsis thaliana]
NCBI Description
                   3358
Seq. No.
Contig ID
                   2370 1.R1040
5'-most EST
                   LIB3073-010-Q1-K1-C5
                   BLASTX
Method
NCBI GI
                   g4097547
BLAST score
                   259
E value
                   3.0e-22
                   110
Match length
```

683



2370 2.R1040 Contig ID 5'-most EST leu701145557.h1 BLASTX Method

g4097547 NCBI GI BLAST score 219 E value 1.0e-17 Match length 118 27

% identity NCBI Description (U64906) ATFP3 [Arabidopsis thaliana]

Seq. No.

2370 4.R1040 Contig ID leu701151209.h1 5'-most EST

3360

3361

Seq. No.

2370 5.R1040 Contig ID 5'-most EST jex700906674.h1

Seq. No.

3362

2371 1.R1040 Contig ID

5'-most EST uC-gmronoir008g01b1

Seq. No.

3363

3364

2372 1.R1040 Contig ID

5'-most EST jC-gmf102220125f01a1

Seq. No.

Contig ID 2379 1.R1040

5'-most EST jC-gmf102220097e06a1

BLASTX Method q3036810 NCBI GI 372 BLAST score E value 4.0e-35 104 Match length 67 % identity

NCBI Description

(AL022373) putative Myc-type transcription factor [Arabidopsis thaliana]

Seq. No.

3365

Contig ID 2379 2.R1040

5'-most EST LIB3109-009-Q1-K1-G5

BLASTX Method NCBI GI q3036810 BLAST score 171 E value 2.0e-12 Match length 48 % identity 67

(AL022373) putative Myc-type transcription factor NCBI Description

[Arabidopsis thaliana]

Seq. No. Contig ID 3366

2379 3.R1040 5'-most EST hrw701060875.hl

Seq. No.

3367

Contig ID

2381 1.R1040

5'-most EST

uC-gmropic097c12b1



Method BLASTX
NCBI GI g4544443
BLAST score 1076
E value 1.0e-117
Match length 317
% identity 68

NCBI Description (AC006592) putative mitochondrial uncoupling protein

[Arabidopsis thaliana]

Seq. No. 3368

Contig ID 2381_2.R1040 5'-most EST uC-gmropic039b05b1

Method BLASTX
NCBI GI g4544443
BLAST score 365
E value 9.0e-35
Match length 111
% identity 70

NCBI Description (AC006592) putative mitochondrial uncoupling protein

[Arabidopsis thaliana]

Seq. No. 3369

Contig ID 2381_3.R1040

5'-most EST LIB3051-053-Q1-K2-B11

Method BLASTX
NCBI GI g4544443
BLAST score 385
E value 4.0e-37
Match length 141
% identity 35

NCBI Description (AC006592) putative mitochondrial uncoupling protein

[Arabidopsis thaliana]

Seq. No. 3370

Contig ID 2381_4.R1040

5'-most EST LIB3106-061-Q1-K1-H10

Seq. No. 3371

Contig ID 2382_1.R1040 5'-most EST leu701147626.h1

Seq. No. 3372

Contig ID 2384_1.R1040

5'-most EST uC-gmflminsoy080g10b1

Method BLASTN
NCBI GI g3821780
BLAST score 35
E value 4.0e-10
Match length 36
% identity 60

NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No. 3373

Contig ID 2388_1.R1040 5'-most EST awf700836736.h1

Method BLASTX NCBI GI g3522938

```
379
BLAST score
                   1.0e-35
E value
                  352
Match length
                   30
% identity
                   (AC004411) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3374
Seq. No.
                   2388 2.R1040
Contig ID
                   jC-gmle01810071h10a1
5'-most EST
                   3375
Seq. No.
                   2388 3.R1040
Contig ID
                   uC-gmrominsoy202c09b1
5'-most EST
                   3376
Seq. No.
                   2390 1.R1040
Contig ID
                   zzp700830963.h1
5'-most EST
                   BLASTN
Method
                   g2980757
NCBI GI
                   51
BLAST score
                   1.0e-19
E value
Match length
                   229
                   87
% identity
                   Arabidopsis thaliana DNA chromosome 4, BAC clone F6I18
NCBI Description
                   (ESSAII project)
                   3377
Seq. No.
                   2392 1.R1040
Contig ID
                   bth700843551.h1
5'-most EST
                   3378
Seq. No.
                   2394 1.R1040
Contig ID
                   LIB3039-005-Q1-E1-E2
5'-most EST
```

Method BLASTX
NCBI GI g100196
BLAST score 1174
E value 1.0e-129
Match length 249
% identity 86

NCBI Description chlorophyll a/b-binding protein (cab-11) - tomato

Seq. No. 3379

Contig ID 2396_1.R1040 5'-most EST ncj700983176.h1

Method BLASTX
NCBI GI g549060
BLAST score 1088
E value 1.0e-136
Match length 363
% identity 69

NCBI Description T-COMPLEX PROTEIN 1, ETA SUBUNIT (TCP-1-ETA) (CCT-ETA) >gi 631656_pir S43058 CCTeta protein eta chain - mouse

>gi_631656_pir__543058 CCTeta protein eta charn = mouse >gi_468504_emb_CAA83274_ (Z31399) CCTeta, eta subunit of the chaperonin containing TCP-1 (CCT) [Mus musculus]

Seq. No. 3380

Contig ID 2397 1.R1040



5'-most EST leu701150938.h1

3381 Seq. No.

2397 2.R1040 Contig ID 5'-most EST $k117\overline{0}1203629.h2$

Method BLASTN NCBI GI g3046847 BLAST score 36 E value 1.0e-10 Match length 104 % identity 84

Arabidopsis thaliana genomic DNA, chromosome5, TAC clone: NCBI Description

K11J9, complete sequence [Arabidopsis thaliana]

Seq. No. 3382

2397 3.R1040 Contig ID 5'-most EST eep700868680.h1

Method BLASTN NCBI GI g3046847 36 BLAST score E value 1.0e-10 Match length 104 % identity 84

Arabidopsis thaliana genomic DNA, chromosome5, TAC clone: NCBI Description

K11J9, complete sequence [Arabidopsis thaliana]

Seq. No. 3383

Contig ID 2400 1.R1040

5'-most EST LIB3051-054-Q1-K2-B11

Method BLASTX NCBI GI q416564 BLAST score 631 E value 9.0e-66 Match length 203 % identity 64

ABCISIC ACID-INDUCIBLE PROTEIN KINASE NCBI Description

>gi_422013_pir__A46408 abscisic acid-inducible

serine/threonine protein kinase homolog - wheat (fragment) >gi 170664 (M94726) protein kinase [Triticum aestivum]

Seq. No.

2403 1.R1040 Contig ID

3384 5'-most EST uC-gmflminsoy045d06b1

Method BLASTN NCBI GI g3449327 BLAST score 35 2.0e-09 E value Match length 168 % identity 85

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MCA23, complete sequence [Arabidopsis thaliana]

Seq. No. 3385

Contig ID 2403 2.R1040

5'-most EST jC-qmst02400007d06a1

Seq. No. 3386



Contig ID 2403_3.R1040

5'-most EST uC-gmrominsoy091c02b1

Seq. No. 338

Contig ID 2403 5.R1040

5'-most EST jC-gmst02400041e01a1

Seq. No. 3388

Contig ID 2404_1.R1040 5'-most EST eep700865952.h1

Seq. No. 3389

Contig ID 2404 2.R1040

5'-most EST uC-gmflminsoy056c06b1

Method BLASTX
NCBI GI g2827661
BLAST score 468
E value 2.0e-46
Match length 230
% identity 43

NCBI Description (AL021637) hyuC-like protein [Arabidopsis thaliana]

Seq. No.

Contig ID 2404 3.R1040

5'-most EST LIB3073-018-Q1-K1-E12

3390

Seq. No. 3391

Contig ID 2404 4.R1040

5'-most EST LIB3051-095-Q1-K1-F9

Method BLASTN
NCBI GI g20740
BLAST score 49
E value 3.0e-18
Match length 187
% identity 86

NCBI Description Pisum sativum mRNA for P protein, a part of glycine

cleavage complex

Seq. No. 3392

Contig ID 2404_5.R1040 5'-most EST zhf700962751.h1

Seq. No. 3393

Contig ID 2405_1.R1040 5'-most EST vzy700753858.h1

Method BLASTX
NCBI GI g1256259
BLAST score 742
E value 1.0e-78
Match length 184
% identity 77

NCBI Description (U50900) voltage-dependent anion channel protein [Spinacia

oleracea]

Seq. No. Contig ID

3394

Contig ID 2406_1.R1040

5'-most EST LIB3139-083-P1-N1-F2



Method BLASTX
NCBI GI g3193292
BLAST score 1576
E value 1.0e-176
Match length 371
% identity 84

NCBI Description (AF069298) similar to ATPases associated with various

cellular activites (Pfam: AAA.hmm, score: 230.91)

[Arabidopsis thaliana]

Seq. No. 3395

Contig ID 2407 1.R1040

5'-most EST LIB3028-045-Q1-B1-A1

Method BLASTX
NCBI GI g4539351
BLAST score 345
E value 1.0e-32
Match length 139
% identity 47

NCBI Description (AL035539) putative protein [Arabidopsis thaliana]

Seq. No. 3396

Contig ID 2410 1.R1040

5'-most EST LIB3028-044-Q1-B1-E3

Method BLASTX
NCBI GI g115833
BLAST score 1118
E value 1.0e-122
Match length 259
% identity 81

NCBI Description CHLOROPHYLL A-B BINDING PROTEIN CP24 10A PRECURSOR

(CAB-10A) (LHCP) >gi_100195_pir__S11877 chlorophyll a/b-binding protein Cab10A - tomato >gi_170394 (M32605)

a-binding protein [Lycopersicon esculentum]

Seq. No. 3397

Contig ID 2410_2.R1040 5'-most EST gsv701045116.h1

Method BLASTX
NCBI GI g115833
BLAST score 1142
E value 1.0e-125
Match length 256
% identity 81

NCBI Description CHLOROPHYLL A-B BINDING PROTEIN CP24 10A PRECURSOR

(CAB-10A) (LHCP) >gi_100195_pir__S11877 chlorophyll a/b-binding protein Cab10A - tomato >gi_170394 (M32605)

a-binding protein [Lycopersicon esculentum]

Seq. No. 3398

Contig ID 2413 1.R1040

5'-most EST LIB3055-005-Q1-N1-H2

Method BLASTX
NCBI GI g3068705
BLAST score 612
E value 3.0e-63
Match length 245

NCBI Description



```
% identity
                  (AF049236) unknown [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  2413 2.R1040
Contig ID
5'-most EST
                  fC-gmro700845160a3
Method
                  BLASTX
NCBI GI
                  q1617200
                  798
BLAST score
                  6.0e-85
E value
Match length
                  161
% identity
                   89
NCBI Description
                  (Y08607) Shaggy-like kinase 6 [Nicotiana tabacum]
Seq. No.
                   3400
Contig ID
                  2413 3.R1040
5'-most EST
                  jC-gmle01810025b09d1
Seq. No.
                   3401
                   2413 4.R1040
Contig ID
5'-most EST
                  fC-qmse700669136r4
Method
                  BLASTX
                  q2598603
NCBI GI
                   173
BLAST score
                   3.0e-12
E value
Match length
                   37
% identity
                  (AJ002315) shaggy-like kinase 59 [Nicotiana tabacum]
NCBI Description
Seq. No.
Contig ID
                   2414 1.R1040
5'-most EST
                   LIB3039-010-Q1-E1-C7
Method
                   BLASTX
NCBI GI
                   q1086252
BLAST score
                   482
E value
                   2.0e-48
Match length
                   142
% identity
                   59
                   sucrose cleavage protein - Potato >gi_707001_bbs_157931
NCBI Description
                   (S74161) sucrolytic enzyme/ferredoxin homolog [Solanum
                   tuberosum=potatoes, cv. Cara, leaf, Peptide, 322 aa]
                   [Solanum tuberosum]
                   3403
Seq. No.
                   2414 2.R1040
Contig ID
                   LIB3092-037-Q1-K1-A2
5'-most EST
                   BLASTX
Method
                   g1086252
NCBI GI
BLAST score
                   209
                   1.0e-16
E value
                   74
Match length
% identity
                   51
```

690

[Solanum tuberosum]

sucrose cleavage protein - Potato >gi_707001_bbs_157931

(S74161) sucrolytic enzyme/ferredoxin homolog [Solanum tuberosum=potatoes, cv. Cara, leaf, Peptide, 322 aa]



Seq. No. 3404

Contig ID 2414 3.R1040

5'-most EST LIB3051-104-Q1-K1-D6

Method BLASTX
NCBI GI g1086252
BLAST score 271
E value 8.0e-24
Match length 95

% identity 52
NCBI Description sucrose cleavage protein - Potato >gi 707001 bbs 157931

(S74161) sucrolytic enzyme/ferredoxin homolog [Solanum tuberosum=potatoes, cv. Cara, leaf, Peptide, 322 aa]

[Solanum tuberosum]

Seq. No. 3405

Contig ID 2418_1.R1040

5'-most EST LIB3028-044-Q1-B1-F3

Seq. No. 3406

Contig ID 2420_1.R1040

5'-most EST LIB3028-044-Q1-B1-F5

Method BLASTX
NCBI GI g1931654
BLAST score 159
E value 9.0e-11
Match length 64
% identity 50

NCBI Description (U95973) BRCA1-associated RING domain protein isolog

[Arabidopsis thaliana]

Seq. No. 3407

Contig ID 2421 1.R1040

5'-most EST LIB3138-034-Q1-N1-C3

Method BLASTX
NCBI GI g4220531
BLAST score 254
E value 1.0e-21
Match length 129
% identity 43

NCBI Description (AL035356) hypothetical protein [Arabidopsis thaliana]

Seq. No. 3408

Contig ID 2421 2.R1040 5'-most EST kl1701214230.h1

Method BLASTX
NCBI GI 94220531
BLAST score 177
E value 1.0e-12
Match length 88
% identity 48

* identity 48
NCBI Description (AL035356) hypothetical protein [Arabidopsis thaliana]

Seq. No. 3409

Contig ID 2422_1.R1040

5'-most EST LIB3040-039-Q1-E1-F10

Method BLASTX NCBI GI g131390



BLAST score 1010 E value 1.0e-110 Match length 243 % identity 79

NCBI Description OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (OEE2) (23 KD SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II)

>gi_280396_pir__JS0771 photosystem II oxygen-evolving

complex protein 2 precursor - garden pea

>qi 20617 emb CAA33557 (X15552) precursor polypeptide (AA

-73 to 186) [Pisum sativum] >gi_344006_dbj_BAA02553_ (D13296) precursor for 23-kDa protein of photosystem II

[Pisum sativum]

Seq. No. 3410

Contig ID 2422 2.R1040

5'-most EST LIB3167-050-P1-K1-F8

Method BLASTN
NCBI GI g4185595
BLAST score 254
E value 1.0e-140
Match length 728
% identity 85

NCBI Description Pisum sativum mRNA for precursor for 23-kDa protein of

photosystem II, complete cds

Seq. No. 3411

Contig ID 2422 3.R1040

5'-most EST LIB3106-095-Q1-K1-F3

Method BLASTX
NCBI GI g131390
BLAST score 474
E value 1.0e-47
Match length 128
% identity 73

NCBI Description

OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (OEE2) (23 KD SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II)

SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II) >gi_280396_pir__JS0771 photosystem II oxygen-evolving

complex protein 2 precursor - garden pea

>gi 20617 emb CAA33557 (X15552) precursor polypeptide (AA

-73 to 186) [Pisum sativum] >gi_344006_dbj_BAA02553_ (D13296) precursor for 23-kDa protein of photosystem II

[Pisum sativum]

Seq. No. 3412

Contig ID 2422 4.R1040

5'-most EST LIB3049-042-Q1-E1-C9

Method BLASTX
NCBI GI g131393
BLAST score 304
E value 9.0e-30
Match length 100
% identity 71

NCBI Description OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (OEE2) (23 KD

SUBUNIT OF OXYGEN-EVOLVING SYSTEM OF PHOTOSYSTEM II) >gi_100360_pir__S15005 photosystem II oxygen-evolving

complex protein 2 - common tobacco

>gi 19911 emb CAA39039.1 (X55354) photosystem II 23kDa



polypeptide [Nicotiana tabacum]

```
Seq. No.
                   3413
                   2423 1.R1040
Contig ID
                   uC\hbox{-} gmflminsoy 089a10b1\\
5'-most EST
                   BLASTX
Method
                   g1402912
NCBI GI
BLAST score
                   646
E value
                   2.0e-67
Match length
                   169
% identity
                   76
                   (X98317) peroxidase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2423 2.R1040
Contig ID
                   jC-gmf102220062h05a1
5'-most EST
Method
                   BLASTX
                   g1402912
NCBI GI
BLAST score
                   216
                   1.0e-26
E value
Match length
                   88
                   73
% identity
NCBI Description (X98317) peroxidase [Arabidopsis thaliana]
                   3415
Seq. No.
                   2423 3.R1040
Contig ID
                   LIB3028-001-Q1-B1-H7
5'-most EST
                   BLASTX
Method
                   g1403138
NCBI GI
                   705
BLAST score
                   5.0e-74
E value
                   180
Match length
                   72
% identity
                   (X98190) peroxidase ATP2a [Arabidopsis thaliana]
NCBI Description
                   >qi 4371288 gb AAD18146 (AC006260) putative peroxidase
                   ATPZa [Arabidopsis thaliana]
                   3416
Seq. No.
                   2423 4.R1040
Contig ID
                   LIB3028-004-Q1-B1-H7
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1402912
BLAST score
                   146
E value
                    4.0e-09
                    33
Match length
                   79
 % identity
                   (X98317) peroxidase [Arabidopsis thaliana]
NCBI Description
 Seq. No.
                    3417
                    2426 1.R1040
 Contig ID
                    uC-gmropic033g08b1
 5'-most EST
                    BLASTN
 Method
 NCBI GI
                    g20899
 BLAST score
                    261
                    1.0e-144
 E value
```

509

89

Match length

% identity



NCBI Description Pea mRNA for Cu/Zn superoxide dismutase II (SOD9)

Seq. No. 3418 Contig ID 2428_1.R1040 5'-most EST wvk700683015.h1

Method BLASTX
NCBI GI g2499614
BLAST score 1423
E value 1.0e-158
Match length 302

% identity 88

NCBI Description MITOGEN-ACTIVATED PROTEIN KINASE HOMOLOG NTF3 (P43)

>gi_481830_pir__S39559 mitogen-activated protein kinase 3
homolog ntf3 - common tobacco >gi_406751_emb_CAA49592_

(X69971) NTF3 [Nicotiana tabacum]

Seq. No. 3419

Contig ID 2428_2.R1040

5'-most EST LIB3049-045-Q1-E1-H6

Method BLASTX
NCBI GI 92499612
BLAST score 187
E value 2.0e-18
Match length 55
% identity 89

NCBI Description MITOGEN-ACTIVATED PROTEIN KINASE HOMOLOG 1 (PMEK1)

>gi_1076650_pir__S52989 mitogen-activated,

extracelluar-regulated protein kinase 1 (EC 2.7.1.-) - garden petunia >gi_603871_emb_CAA58466_ (X83440) MAP/ERK

kinase 1 [Petunia \bar{x} hybrida]

Seq. No. 3420

Contig ID 2429_1.R1040

5'-most EST LIB3106-056-Q1-K1-H1

Seq. No. 3421

Contig ID 2429_2.R1040
5'-most EST g5688463
Method BLASTX
NCBI GI g1582580
BLAST score 1759
E value 0.0e+00
Match length 364
% identity 90

% identity 90 NCBI Description caffeic acid O-methyltransferase [Stylosanthes humilis]

Seq. No. 3422

Contig ID 2429_3.R1040

5'-most EST LIB3139-024-P1-N1-A2

Method BLASTN
NCBI GI 94468045
BLAST score 76
E value 1.0e-34
Match length 156
% identity 87

% identity 87
NCBI Description V.planifolia mRNA for methyltransferase



```
3423
Seq. No.
                   2429 5.R1040
Contig ID
                   ssr7\overline{0}0554583.h1
5'-most EST
                   BLASTN
Method
                   g258157
NCBI GI
                   53
BLAST score
                   7.0e-21
E value
Match length
                   81
% identity
                   91
                   psbD=reaction center protein D2 {5'region} [peas,
NCBI Description
                   Chloroplast, 90 nt]
                   3424
Seq. No.
                   2431 1.R1040
Contig ID
                   zhf7\overline{0}0961150.h1
5'-most EST
                   BLASTX
Method
                   g2829898
NCBI GI
BLAST score
                   285
                   2.0e-25
E value
                   89
Match length
                   61
% identity
                  (AC002311) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3425
Seq. No.
                   2432 1.R1040
Contig ID
                   LIB3028-044-Q1-B1-C2
5'-most EST
                   3426
Seq. No.
                   2433 1.R1040
Contig ID
                   LIB3028-040-Q1-B1-E11
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2642153
                   141
BLAST score
E value
                   1.0e-08
                   89
Match length
% identity
NCBI Description
                   (AC003000) unknown protein [Arabidopsis thaliana]
                   >gi 2795810 (AC003674) unknown protein [Arabidopsis
                   thaliana]
                   3427
Seq. No.
Contig ID
                   2436 1.R1040
5'-most EST
                   LIB3052-011-Q1-N1-A2
Method
                   BLASTX
                   q2982465
NCBI GI
BLAST score
                   496
                   9.0e-50
E value
                   203
Match length
% identity
                   50
                   (AL022223) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3428
Contig ID
                   2440 1.R1040
                   qsf700698384.hl
5'-most EST
```

BLASTX

1981

q510876

Method

NCBI GI BLAST score



```
0.0e+00
E value
Match length
                   449
% identity
                   (X80051) NADP dependent malic enzyme [Phaseolus vulgaris]
NCBI Description
Seq. No.
                   2440 2.R1040
Contig ID
                   eep700869552.hl
5'-most EST
Method
                   BLASTX
                   g1561774
NCBI GI
BLAST score
                   733
                   7.0e-78
E value
                   188
Match length
                   81
% identity
NCBI Description (U67426) malate dehydrogenase [Vitis vinifera]
                   3430
Seq. No.
                   2440 5.R1040
Contig ID
                   k117\overline{0}1206630.h1
5'-most EST
Method
                   BLASTX
                   g1708924
NCBI GI
                   254
BLAST score
E value
                   6.0e-22
Match length
                   86
                   85
% identity
                   MALATE OXIDOREDUCTASE (MALIC ENZYME) (ME) (NADP-DEPENDENT
NCBI Description
                   MALIC ENZYME) (NADP-ME) >gi_515759 (L34836) malate
                   dehydrogenase (NADP+) [Vitis vinifera]
                   3431
Seq. No.
                   2440 6.R1040
Contig ID
                   crh7\overline{0}0854758.h1
5'-most EST
                   3432
Seq. No.
                   2442 1.R1040
Contig ID
                   q4293373
5'-most EST
                   BLASTX
Method
                   g585264
NCBI GI
                   374
BLAST score
                   1.0e-35
E value
                   185
Match length
                   42
 % identity
                   HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE (HGPRT)
NCBI Description
                    (HGPRTASE) >gi_2127112_pir__S66098 hypoxanthine-guanine
                   phosphoribosyltransferase - Bacillus subtilis
                   >gi 467457 dbj BAA05303 (D26185) hypoxanthine-guanine
                   phosphoribosyltransferase [Bacillus subtilis]
                   >gi_2632335_emb_CAB11844_ (Z99104) hypoxanthine-guanine
                   phosphoribosyltransferase [Bacillus subtilis]
                    3433
 Seq. No.
                    2443 1.R1040
 Contig ID
```

 Seq. No.
 3433

 Contig ID
 2443_1.R1040

 5'-most EST
 gsv701049235.h1

 Method
 BLASTX

 NCBI GI
 g1903364

 BLAST score
 97

E value

1.0e-08



112 Match length 34 % identity (AC000104) EST gb_T45093 comes from this gene. [Arabidopsis NCBI Description thaliana] 3434 Seq. No. 2443 2.R1040 Contig ID LIB3170-062-Q1-J1-H8 5'-most EST Seq. No. 3435 2444 1.R1040 Contig ID uC-gmropic005h11b1 5'-most EST BLASTX Method NCBI GI q1352088 1834 BLAST score 0.0e+00E value 466 Match length 73 % identity CITRATE SYNTHASE, MITOCHONDRIAL PRECURSOR >gi 624676 NCBI Description (U19481) citrate synthase precursor [Citrus maxima] Seq. No. 3436 2444 2.R1040 Contig ID vwf700674464.h1 5'-most EST Method BLASTX q3493367 NCBI GI 181 BLAST score 2.0e-13 E value 53 Match length % identity 68 (AB017159) citrate synthase [Daucus carota] NCBI Description 3437 Seq. No. 2447 1.R1040 Contig ID jC-gmro02910051g10a1 5'-most EST BLASTX Method g2118017 NCBI GI BLAST score 639 8.0e-67 E value 131 Match length 81 % identity non-S-RNase (EC 3.1.-.-) - Japanese pear NCBI Description >gi 1526417 dbj BAA08475_ (D49529) ribonuclease [Pyrus pyrifolia] 3438 Seq. No. 2448 1.R1040 Contig ID LIB3167-004-P1-K1-D7 5'-most EST Method BLASTX g399082 NCBI GI BLAST score 676

6.0e-71E value 248 Match length % identity 60

ATP SYNTHASE DELTA CHAIN, CHLOROPLAST PRECURSOR NCBI Description

>gi_322713_pir__S28171 H+-transporting ATP synthase (EC 3.6.1.34) delta chain, chloroplast - garden pea >gi_169045



(M94558) ATP synthase delta subunit [Pisum sativum]

3439 Seq. No. 2448 2.R1040 Contig ID LIB3138-032-Q1-N1-D4 5'-most EST 3440 Seq. No. 2448 3.R1040 Contig ID rlr700899859.h1 5'-most EST BLASTX Method g399082 NCBI GI 157 BLAST score 2.0e-10 E value 116 Match length 36 % identity ATP SYNTHASE DELTA CHAIN, CHLOROPLAST PRECURSOR NCBI Description >gi 322713 pir S28171 H+-transporting ATP synthase (EC 3.6.1.34) delta chain, chloroplast - garden pea >gi_169045 (M94558) ATP synthase delta subunit [Pisum sativum] 3441 Seq. No. 2450 1.R1040 Contig ID LIB3049-038-Q1-E1-B10 5'-most EST BLASTX Method q643469 NCBI GI 798 BLAST score 4.0e-85 E value 252 Match length 62 % identity NCBI Description (U19886) unknown [Lycopersicon esculentum] 3442 Seq. No. 2450 2.R1040 Contig ID LIB3106-037-Q1-K1-H6 5'-most EST BLASTX Method q3158376 NCBI GI 298 BLAST score E value 7.0e-27 114 Match length 56 % identity (AF035385) unknown [Arabidopsis thaliana] NCBI Description Seq. No. 3443 2450 3.R1040 Contig ID LIB3170-031-Q1-J1-A6 5'-most EST BLASTX Method q643469 NCBI GI BLAST score 323 8.0e-30 E value Match length 82 71 % identity (U19886) unknown [Lycopersicon esculentum] NCBI Description

Seq. No. 3444

Contig ID 2450_4.R1040 5'-most EST gsv701045693.h1

Method BLASTX

BLAST score

E value Match length 57 2.0e-23

101

```
g3434973
NCBI GI
                   251
BLAST score
                   5.0e-21
E value
Match length
                   66
                   73
% identity
                   (AB008106) ethylene responsive element binding factor 4
NCBI Description
                   [Arabidopsis thaliana]
                   3445
Seq. No.
                   2450 5.R1040
Contig ID
                   wrg700790552.h2
5'-most EST
                   BLASTN
Method
                   q4099920
NCBI GI
                   63
BLAST score
                   2.0e-26
E value
                   143
Match length
                   90
% identity
                   Stylosanthes hamata EREBP-3 homolog mRNA, complete cds
NCBI Description
                   3446
Seq. No.
                   2450 6.R1040
Contig ID
                   LIB3170-031-Q1-K1-A12
5'-most EST
                   BLASTX
Method
NCBI GI
                   g643469
                   264
BLAST score
                   1.0e-22
E value
Match length
                   71
                   72
% identity
NCBI Description (U19886) unknown [Lycopersicon esculentum]
Seq. No.
                   3447
                   2450 10.R1040
Contig ID
5'-most EST
                   jC-gmst02400057b12d1
                   3448
Seq. No.
Contig ID
                   2450 11.R1040
5'-most EST
                   jC-gmle01810042e11a1
Method
                   BLASTX
                   g4099921
NCBI GI
BLAST score
                   287
                   1.0e-25
E value
Match length
                   108
% identity
                   (U91982) EREBP-3 homolog [Stylosanthes hamata]
NCBI Description
                   3449
Seq. No.
                   2450 12.R1040
Contig ID
                   epx701107465.h1
5'-most EST
                   3450
Seq. No.
                   2450 14.R1040
Contig ID
5'-most EST
                   dpv7\overline{0}1101450.h1
Method
                   BLASTN
                   g3434972
NCBI GI
```



% identity 89
NCBI Description Arabidopsis thaliana AtERF-4 mRNA for ethylene responsive element binding factor 4, complete cds

Seq. No. 3451

Contig ID 2453_1.R1040

5'-most EST LIB3030-006-Q1-B1-D8

Method BLASTX
NCBI GI 94105097
BLAST score 281
E value 4.0e-41
Match length 110
% identity 80

NCBI Description (AF043255) MADS box protein 26 [Cucumis sativus]

Seq. No. 3452

Contig ID 2454_1.R1040

5'-most EST LIB3109-010-Q1-K1-E1

Method BLASTX
NCBI GI g2570047
BLAST score 339
E value 5.0e-31
Match length 164
% identity 43

NCBI Description (Y09234) MSTK2S kinase-like protein [Mus musculus]

Seq. No. 3453

Contig ID 2457 1.R1040

5'-most EST LIB3030-008-Q1-B1-E2

Method BLASTX
NCBI GI g3386614
BLAST score 415
E value 1.0e-40
Match length 115
% identity 40

NCBI Description (AC004665) putative transcription factor SF3 [Arabidopsis

thaliana]

Seq. No. 3454

Contig ID 2458_1.R1040

5'-most EST LIB3028-044-Q1-B1-B11

Method BLASTX
NCBI GI g3123059
BLAST score 185
E value 5.0e-13
Match length 180
% identity 31

NCBI Description HYPOTHETICAL PROTEIN KIAA0253 >gi_1665773_dbj_BAA13383_

(D87442) KIAA0253 [Homo sapiens]

Seq. No. 3455

Contig ID 2459_1.R1040

5'-most EST LIB3028-043-Q1-B1-H5

Method BLASTX
NCBI GI g2264373
BLAST score 335
E value 4.0e-31



```
Match length 169 % identity 43
```

NCBI Description (AC002354) putative NAM/no apical meristem protein

[Arabidopsis thaliana]

Seq. No. 3456

Contig ID 2461_1.R1040

5'-most EST jC-gmle01810068f07a1

Method BLASTX
NCBI GI g1749474
BLAST score 151
E value 2.0e-09
Match length 96
% identity 40

NCBI Description (D89133) similar to Saccharomyces cerevisiae Lph16p,

GENBANK Accession Number U43503 [Schizosaccharomyces pombe]

Seq. No. 3457

Contig ID 2462_1.R1040 5'-most EST leu701151985.h1

Method BLASTX
NCBI GI g3747111
BLAST score 627
E value 3.0e-65
Match length 211
% identity 55

NCBI Description (AF095641) MTN3 homolog [Arabidopsis thaliana]

Seq. No. 3458

Contig ID 2462_2.R1040 5'-most EST epx701105770.h1

Method BLASTN
NCBI GI g11565
BLAST score 308
E value 1.0e-173
Match length 316
% identity 99

NCBI Description Soybean chloroplast psb A gene coding for photosystem II

thylakoid membrane protein

Seq. No. 3459

Contig ID 2465_1.R1040 5'-most EST bth700845088.h1

Method BLASTX
NCBI GI g2792297
BLAST score 334
E value 8.0e-31
Match length 66
% identity 83

NCBI Description (AF039183) GAST-like gene product [Fragaria x ananassa]

Seq. No. 3460

Contig ID 2465_2.R1040

5'-most EST LIB3093-015-Q1-K1-A3

Method BLASTX NCBI GI g2792297 BLAST score 334



```
E value
                   5.0e - 31
Match length
                   66
% identity
                   83
NCBI Description
                  (AF039183) GAST-like gene product [Fragaria x ananassa]
Seq. No.
                   3461
Contig ID
                   2465 3.R1040
5'-most EST
                  LIB3107-060-Q1-K1-G3
Method
                   BLASTX
                   g2792297
```

NCBI GI BLAST score 342 E value 8.0e-32 Match length 66 % identity 86

NCBI Description (AF039183) GAST-like gene product [Fragaria x ananassa]

3462 Seq. No. Contig ID 2466 1.R1040 5'-most EST $jsh7\overline{0}1067168.h1$ Method BLASTX NCBI GI g4572679

BLAST score 322 E value 2.0e-29 Match length 69 % identity 87

NCBI Description (AC006954) RSZp22 splicing factor; contains RNA recognition

motif [Arabidopsis thaliana]

Seq. No. 3463

Contig ID 2466 2.R1040

5'-most EST LIB3051-005-Q1-E1-D11

Method BLASTX NCBI GI g4572679 BLAST score 406 E value 2.0e-39 Match length 104 % identity 76

(AC006954) RSZp22 splicing factor; contains RNA recognition NCBI Description

motif [Arabidopsis thaliana]

Seq. No.

2466 3.R1040 Contig ID

5'-most EST jC-gmf102220115e04a1

3464

Seq. No.

3465

Contig ID 2466 4.R1040

5'-most EST jC-gmst02400026d01a1

Method BLASTX NCBI GI g4572679 BLAST score 278 E value 1.0e-24 Match length 60 % identity 87

NCBI Description (AC006954) RSZp22 splicing factor; contains RNA recognition

motif [Arabidopsis thaliana]

Seq. No. 3466



```
Contig ID
                   2466 5.R1040
5'-most EST
                  LIB3106-040-Q1-K1-B2
Seq. No.
                   3467
Contig ID
                   2467 1.R1040
5'-most EST
                  LIB3109-028-Q1-K1-F3
Method
                  BLASTX
NCBI GI
                  q3128228
BLAST score
                  871
E value
                  9.0e-94
Match length
                  178
% identity
                  90
NCBI Description
```

3468

(AC004077) putative ribosomal protein L18A [Arabidopsis thaliana] >gi_3337376 (AC004481) putative ribosomal protein L18A [Arabidopsis thaliana]

Contig ID 2467 2.R1040 5'-most EST LIB3028-043-Q1-B1-G5 Method BLASTX NCBI GI g3128228 BLAST score 601 E value 1.0e-62 Match length 123 % identity 91

Seq. No.

NCBI Description (AC004077) putative ribosomal protein L18A [Arabidopsis thaliana] >gi 3337376 (AC004481) putative ribosomal protein

L18A [Arabidopsis thaliana]

Seq. No. 3469 Contig ID 2469 1.R1040 5'-most EST $leu7\overline{0}1146773.h1$ Method BLASTX NCBI GI g4220524 BLAST score 185

E value 2.0e-13 Match length 68 % identity 54

NCBI Description (AL035356) putative protein [Arabidopsis thaliana]

Seq. No. 3470

Contig ID 2469 2.R1040 5'-most EST kll701203514.h2

Seq. No. 3471

Contig ID 2469 3.R1040 5'-most EST pxt700942179.hl

Seq. No. 3472 Contig ID

2472 1.R1040 5'-most EST LIB3028-043-Q1-B1-H1

Method BLASTX NCBI GI q4538897 BLAST score 461 E value 4.0e-46 Match length 134 % identity 63



NCBI Description (AL049482) AX110P-like protein [Arabidopsis thaliana]

Seq. No. 3473

Contig ID 2474_1.R1040

5'-most EST LIB3028-020-Q1-B1-G11

Seq. No. 3474

Contig ID 2474 3.R1040

5'-most EST jC-gmst02400036d12d2

Seq. No. 3475

Contig ID 2477 1.R1040

5'-most EST LIB3170-053-Q1-K1-A2

Method BLASTN
NCBI GI g1142620
BLAST score 296
E value 1.0e-165
Match length 1181
% identity 86

NCBI Description Phaseolus vulgaris phaseolin G-box binding protein PG2

(PG2) mRNA, partial cds

Seq. No. 3476

Contig ID 2477_2.R1040

5'-most EST uC-gmropic075f09b1

Method BLASTN
NCBI GI g1142620
BLAST score 208
E value 1.0e-113
Match length 390
% identity 88

NCBI Description Phaseolus vulgaris phaseolin G-box binding protein PG2

(PG2) mRNA, partial cds

Seq. No. 3477

Contig ID 2478_1.R1040

5'-most EST LIB3051-112-Q1-K1-E5

Method BLASTX
NCBI GI g3522945
BLAST score 223
E value 3.0e-18
Match length 130
% identity 34

NCBI Description (AC004411) putative cytochrome P450 [Arabidopsis thaliana]

Seq. No. 3478

Contig ID 2480 1.R1040

5'-most EST fC-gmro700866825a3

Method BLASTX
NCBI GI g3461814
BLAST score 1410
E value 1.0e-156
Match length 397
% identity 67

NCBI Description (AC004138) hypothetical protein [Arabidopsis thaliana]

Seq. No. 3479



```
Contig ID
                   2480 2.R1040
5'-most EST
                   uC-gmrominsoy308c02b1
Seq. No.
                   3480
Contig ID
                   2481 1.R1040
5'-most EST
                   jex700906745.h1
Method
                   BLASTX
NCBI GI
                   g4490318
BLAST score
                   1704
E value
                   0.0e+00
Match length
                   382
% identity
                   85
NCBI Description
                   (AL035678) 2-dehydro-3-deoxyphosphoheptonate aldolase
                   [Arabidopsis thaliana]
                   3481
Seq. No.
Contig ID
                   2481 2.R1040
5'-most EST
                   LIB3029-012-Q1-B1-B5
Method
                   BLASTX
NCBI GI
                   g99743
BLAST score
                   741
E value
                   1.0e-78
Match length
                   156
% identity
                   92
NCBI Description
                   2-dehydro-3-deoxyphosphoheptonate aldolase (EC 4.1.2.15) 2
                   - Arabidopsis thaliana
Seq. No.
                   3482
Contig ID
                   2485_1.R1040
5'-most EST
                   LIB3049-055-Q1-E1-G3
Method
                   BLASTX
NCBI GI
                   g4321188
BLAST score
                   337
E value
                   3.0e-31
Match length
                   146
% identity
                   58
                   (AF021807) low molecular weight heat-shock protein [Corylus
NCBI Description
                  avellana]
Seq. No.
                   3483
                  2485 2.R1040
Contig ID
5'-most EST
                  LIB3053-002-Q1-B1-E1
Method
                  BLASTX
NCBI GI
                  g4321188
BLAST score
                  310
E value
                  3.0e-28
Match length
                  94
% identity
                  65
NCBI Description
                  (AF021807) low molecular weight heat-shock protein [Corylus
                  avellana]
```

Seq. No. 3484

Contig ID 2486 1.R1040 5'-most EST rca701001119.h1 Method BLASTX

Method BLASTX NCBI GI g732003 BLAST score 204



E value 7.0e-16 Match length 153 % identity 35

NCBI Description OUTER MEMBRANE LIPOPROTEIN BLC PRECURSOR

>gi_2125977_pir__184534 outer membrane lipoprotein Escherichia coli >gi_536993 (U14003) ORF_f177 [Escherichia coli] >gi_717134 (U21726) lipocalin precursor [Escherichia coli] >gi_1790592 (AE000487) outer membrane lipoprotein

(lipocalin) [Escherichia coli]

Seq. No. 3485

Contig ID 2487_1.R1040 5'-most EST gsv701056217.h1

Method BLASTX
NCBI GI g4454475
BLAST score 231
E value 4.0e-19
Match length 67
% identity 63

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No. 3486

Contig ID 2488 1.R1040

5'-most EST LIB3138-082-P1-N1-D4

Method BLASTX
NCBI GI g4220481
BLAST score 418
E value 3.0e-43
Match length 160
% identity 59

NCBI Description (AC006069) unknown protein [Arabidopsis thaliana]

Seq. No.

3487

Contig ID 2488 2.R1040

5'-most EST uC-gmrominsoy093c03b1

Method BLASTX
NCBI GI g4508082
BLAST score 363
E value 2.0e-34
Match length 71
% identity 90

NCBI Description (AC005882) Putative RNA polymerase II subunit Rpb10

[Arabidopsis thaliana]

Seq. No. 3488

Contig ID 2488 3.R1040 5'-most EST zhf700962530.h1

Method BLASTX
NCBI GI g4508082
BLAST score 361
E value 3.0e-34
Match length 71
% identity 92

NCBI Description (AC005882) Putative RNA polymerase II subunit Rpb10

[Arabidopsis thaliana]

Seq. No. 3489



```
2488 4.R1040
Contig ID
5'-most EST
                  jC-gmro02800038a07d1
                  3490
Seq. No.
                  2488 6.R1040
Contig ID
5'-most EST
                  eep7\overline{0}0865123.h1
                  BLASTX
Method
NCBI GI
                  q4220481
BLAST score
                  242
E value
                  2.0e-20
Match length
                  85
                  55
% identity
NCBI Description (AC006069) unknown protein [Arabidopsis thaliana]
Seq. No.
                  3491
                  2488 8.R1040
Contig ID
5'-most EST
                  pcp700992848.hl
Seq. No.
                  3492
Contig ID
                  2489 1.R1040
5'-most EST
                  LIB3170-007-Q1-K1-E9
                   3493
Seq. No.
Contig ID
                  2490 1.R1040
5'-most EST
                  LIB3028-043-Q1-B1-F7
Method
                  BLASTX
                  g1353018
NCBI GI
BLAST score
                  196
                  2.0e-14
E value
Match length
                  157
% identity
                  31
NCBI Description HYPOTHETICAL 34.4 KD PROTEIN IN IDS2-MPI2 INTERGENIC REGION
                  >gi 1077841 pir S55168 hypothetical protein YJL145w -
                   yeast (Saccharomyces cerevisiae) >gi_854552_emb_CAA60810_
                   (X87371) ORF10 [Saccharomyces cerevisiae]
                   >gi_1015557_emb_CAA89440_ (Z49420) ORF YJL145w
                   [Saccharomyces cerevisiae]
Seq. No.
                   3494
                   2491 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy119a01b1
                  BLASTX
Method
                   g1684851
NCBI GI
BLAST score
                   326
E value
                  3.0e - 30
Match length
                  117
% identity
NCBI Description (U77935) DnaJ-like protein [Phaseolus vulgaris]
```

Seq. No. 3495

Contig ID 2492_1.R1040 5'-most EST gsv701055713.h1

Seq. No. 3496

Contig ID 2492_2.R1040

5'-most EST LIB3106-072-P1-K1-B6

% identity

58



```
Seq. No.
                   3497
Contig ID
                   2492 3.R1040
5'-most EST
                   LIB3109-027-Q1-K1-F8
                   3498
Seq. No.
                   2494 1.R1040
Contig ID
5'-most EST
                   LIB3170-005-Q1-K1-F3
Method
                   BLASTX
NCBI GI
                   q119958
BLAST score
                   459
E value
                   2.0e-45
Match length
                   152
% identity
                   59
                   FERREDOXIN III PRECURSOR (FD III) >gi_168473 (M73831)
NCBI Description
                   ferredoxin [Zea mays] >gi_1864001_dbj_BAA19251_ (AB001387) Fd III [Zea mays] >gi_444686_prf__1907324C
                   ferredoxin:ISOTYPE=III [Zea mays]
Seq. No.
                   3499
                   2494 2.R1040
Contig ID
                   zhf700964962.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4115534
BLAST score
                   481
                   5.0e-48
E value
Match length
                   226
% identity
NCBI Description
                   (AB012114) UDP-glycose:flavonoid glycosyltransferase [Vigna
                   mungo]
Seq. No.
                   3500
Contig ID
                   2497 1.R1040
5'-most EST
                   jsh701064037.h1
Seq. No.
                   3501
                   2498 1.R1040
Contig ID
                   leu701153865.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4415930
BLAST score
                   207
                   3.0e-16
E value
Match length
                   66
                   58
% identity
NCBI Description
                   (AC006418) unknown protein [Arabidopsis thaliana]
                   >gi 4559389 gb_AAD23049.1_AC006526_14 (AC006526) unknown
                   protein [Arabidopsis thaliana]
                   3502
Seq. No.
Contig ID
                   2498 2.R1040
5'-most EST
                   LIB3049-032-Q1-E1-E1
Method
                   BLASTX
NCBI GI
                   q4415930
BLAST score
                   225
E value
                   1.0e-18
                   69
Match length
```

NCBI Description (AC006418) unknown protein [Arabidopsis thaliana]

Match length

% identity

74

43



>gi_4559389_gb_AAD23049.1_AC006526_14 (AC006526) unknown
protein [Arabidopsis thaliana]

```
3503
Seq. No.
                  2501 1.R1040
Contig ID
                  LIB3028-043-Q1-B1-D7
5'-most EST
Method
                  BLASTX
                  g4006922
NCBI GI
BLAST score
                  158
E value
                  1.0e-10
Match length
                  35
% identity
                  77
                  (Z99708) cytochrome P450 like protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  3504
                  2503 1.R1040
Contig ID
                  uC-gmrominsoy135g07b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2462826
BLAST score
                  759
                  1.0e-109
E value
Match length
                  477
                  45
% identity
NCBI Description (AF000657) unknown protein [Arabidopsis thaliana]
Seq. No.
                  3505
Contig ID
                  2503 2.R1040
5'-most EST
                  zhf700956794.h1
Method
                  BLASTX
NCBI GI
                  q2462826
BLAST score
                  240
                  3.0e-20
E value
Match length
                  99
% identity
                  45
NCBI Description
                  (AF000657) unknown protein [Arabidopsis thaliana]
                   3506
Seq. No.
                   2504 1.R1040
Contig ID
                  LIB3039-018-Q1-E1-E6
5'-most EST
                  BLASTX
Method
                  g3241945
NCBI GI
BLAST score
                   438
E value
                   4.0e-43
Match length
                  146
                  55
% identity
NCBI Description (AC004625) unknown protein [Arabidopsis thaliana]
Seq. No.
                   3507
Contig ID
                  2505 1.R1040
5'-most EST
                  uC-gmflminsoy053h03b1
Method
                  BLASTX
NCBI GI
                  g2935523
BLAST score
                  142
E value
                  1.0e-08
```



NCBI Description (AF049066) 21 kD protein precursor [Pinus radiata]

Seq. No.

3508

Contig ID

2505 2.R1040

5'-most EST

LIB3028-043-Q1-B1-E10

Method NCBI GI BLASTX q112717

BLAST score

235

E value Match length 2.0e-19 128

% identity

40

NCBI Description

21 KD PROTEIN PRECURSOR (1.2 PROTEIN) >gi 82050 pir S10911

hypothetical protein precursor - carrot

>gi_18312_emb_CAA36642_ (X52395) precursor polypeptide (AA

-22 to 171) [Daucus carota]

Seq. No.

3509

Contig ID 5'-most EST 2508 1.R1040 zhf700963215.h1

Method NCBI GI BLASTX

BLAST score

q2252871

E value

162

2.0e-23

Match length

134

% identity

48

NCBI Description (AF013294) No definition line found [Arabidopsis thaliana]

Seq. No.

3510

Contig ID

2510 1.R1040

5'-most EST

LIB3170-036-Q1-K2-A6

Method

BLASTX g2895866

NCBI GI

BLAST score

399

E value

1.0e-38

Match length

154

% identity

59

NCBI Description

(AF045770) methylmalonate semi-aldehyde dehydrogenase

[Oryza sativa]

Seq. No.

3511

Contig ID

2510 2.R1040

5'-most EST

LIB3092-048-Q1-K1-F8

Seq. No.

3512

Contig ID

2511 1.R1040

5'-most EST

LIB3138-006-Q1-N1-C7

Method NCBI GI BLASTX g2529683

BLAST score

426

E value

9.0e-42

Match length

% identity

186 48

NCBI Description (AC002535) unknown protein [Arabidopsis thaliana]

Seq. No.

3513

Contig ID

2512 1.R1040

5'-most EST

LIB3049-046-Q1-E1-H9



Seq. No. 3514

Contig ID 2513_1.R1040

5'-most EST uC-gmropic010e09b1

Method BLASTX
NCBI GI g728882
BLAST score 808
E value 2.0e-86
Match length 178
% identity 87

NCBI Description ADP-RIBOSYLATION FACTOR 3 >gi_541846_pir__S41938

4

ADP-ribosylation factor 3 - Arabidopsis thaliana

>gi_453191_emb_CAA54564_ (X77385) ADP-ribosylation factor 3

[Arabidopsis thaliana]

Seq. No. 3515

Contig ID 2513_2.R1040 5'-most EST leu701145405.h1

Method BLASTN
NCBI GI 9453190
BLAST score 95
E value 8.0e-46
Match length 191
% identity 87

NCBI Description A.thaliana mRNA for ADP-ribosylation factor

Seq. No. 3516

Contig ID 2517_1.R1040 5'-most EST rca701001442.h1

Method BLASTX
NCBI GI g4521322
BLAST score 391
E value 1.0e-63
Match length 279
% identity 50

NCBI Description (U11790) mitotic centromere-associated kinesin [Cricetulus

griseus]

Seq. No.

Contig ID 2518 1.R1040

5'-most EST LIB3028-043-Q1-B1-B8

3517

Method BLASTX
NCBI GI g4490317
BLAST score 239
E value 4.0e-20
Match length 131
% identity 39

NCBI Description (AL035678) putative protein [Arabidopsis thaliana]

Seq. No. 3518

Contig ID 2519_1.R1040 5'-most EST leu701153841.h1

Seq. No. 3519

Contig ID 2523_1.R1040 5'-most EST ssr700556872.h1

Method BLASTX



```
g3123515
NCBI GI
                  599
BLAST score
                  3.0e-62
E value
                  126
Match length
% identity
                  (Y08761) Mago Nashi-like protein [Euphorbia lagascae]
NCBI Description
                  3520
Seq. No.
                  2525 1.R1040
Contig ID
5'-most EST
                  ncj700979085.hl
                  BLASTX
Method
                  g3193292
NCBI GI
                  502
BLAST score
                  7.0e-51
E value
                  159
Match length
                  58
% identity
                  (AF069298) similar to ATPases associated with various
NCBI Description
                  cellular activites (Pfam: AAA.hmm, score: 230.91)
                   [Arabidopsis thaliana]
                   3521
Seq. No.
Contig ID
                   2526 1.R1040
                  pcp700992672.h1
5'-most EST
                   3522
Seq. No.
                   2526 2.R1040
Contig ID
5'-most EST
                  LIB3094-099-Q1-K1-G5
                   3523
Seq. No.
                   2527 1.R1040
Contig ID
                   wrg700789265.h2
5'-most EST
Method
                   BLASTX
                   q2367392
NCBI GI
                   358
BLAST score
                   1.0e-33
E value
                   330
Match length
                   29
% identity
                  (U82513) random slug cDNA25 protein [Dictyostelium
NCBI Description
                   discoideum]
                   3524
Seq. No.
                   2527 2.R1040
Contig ID
                   uC-gmflminsoy011a04b1
5'-most EST
                   BLASTX
Method
                   g2367392
NCBI GI
                   288
BLAST score
                   1.0e-25
E value
                   170
Match length
% identity
                   36
                   (U82513) random slug cDNA25 protein [Dictyostelium
NCBI Description
                   discoideum]
```

 Seq. No.
 3525

 Contig ID
 2528_1.R1040

 5'-most EST
 g4296284

 Method
 BLASTN

NCBI GI g3820639



```
BLAST score 281
E value 1.0e-156
Match length 782
% identity 84
```

NCBI Description C.arietinum mRNA for cysteine proteinase

Seq. No. 3526

Contig ID 2528 2.R1040

5'-most EST LIB3093-045-Q1-K1-B2

Method BLASTN
NCBI GI g3820639
BLAST score 101
E value 5.0e-49
Match length 312
% identity 86

NCBI Description C.arietinum mRNA for cysteine proteinase

Seq. No. 3527

Contig ID 2528 3.R1040

5'-most EST LIB3092-058-Q1-K1-B11

Method BLASTN
NCBI GI g3820639
BLAST score 203
E value 1.0e-110
Match length 632
% identity 84

NCBI Description C.arietinum mRNA for cysteine proteinase

Seq. No. 3528

Contig ID 2528 4.R1040 5'-most EST q5510260 Method BLASTN NCBI GI q3820639 BLAST score 169 6.0e-90 E value Match length 404 86 % identity

NCBI Description C.arietinum mRNA for cysteine proteinase

Seq. No. 3529

Contig ID 2528 5.R1040

5'-most EST jC-gmfl02220070g10a1

Method BLASTX
NCBI GI g1134882
BLAST score 226
E value 2.0e-18
Match length 94
% identity 49

NCBI Description (Z68291) cysteine protease [Pisum sativum]

Seq. No.

3530

Contig ID 2529 1.R1040

5'-most EST LIB3107-062-Q1-K1-D8

Method BLASTN
NCBI GI g1370145
BLAST score 386
E value 0.0e+00



Match length % identity

L.japonicus mRNA for small GTP-binding protein, RAB11C NCBI Description

*

Seq. No.

3531 Contig ID

2529 2.R1040 LIB3107-060-Q1-K1-A7 5'-most EST

Method BLASTN q1370145 NCBI GI 175 BLAST score E value 1.0e-93 267 Match length

91 % identity

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB11C

Seq. No.

3532

Contig ID 2530 1.R1040

5'-most EST LIB3170-052-Q1-J1-B11

3533

Seq. No.

2532 1.R1040 Contig ID

5'-most EST LIB3028-043-Q1-B1-A11

3534 Seq. No.

2534 1.R1040 Contig ID

jC-gmro02910016c11d1 5'-most EST

Method BLASTX g2462825 NCBI GI BLAST score 468 3.0e-46 E value 180 Match length % identity 62

(AF000657) contains Procite 'RNP1' putative RNA-binding NCBI Description

region [Arabidopsis thaliana]

3535 Seq. No.

Contig ID 2534 2.R1040 5'-most EST q5606849

Seq. No. 3536

Contig ID 2534 3.R1040

5'-most EST jC-gmro02910002g10a1

Seq. No. 3537

2534 4.R1040 Contig ID asn701135234.h1 5'-most EST

BLASTX Method g2462825 NCBI GI BLAST score 450 1.0e-44 E value 100 Match length 89 % identity

(AF000657) contains Procite 'RNP1' putative RNA-binding NCBI Description

region [Arabidopsis thaliana]

3538 Seq. No.

Contig ID 2534 5.R1040



```
dpv701101840.h1
5'-most EST
Method
                  BLASTX
                  g2462825
NCBI GI
                  404
BLAST score
E value
                  4.0e-39
                  88
Match length
                  89
% identity
                 (AF000657) contains Procite 'RNP1' putative RNA-binding
NCBI Description
                  region [Arabidopsis thaliana]
Seq. No.
                  3539
                  2534 6.R1040
Contig ID
5'-most EST
                  zsg701124648.hl
Seq. No.
                  3540
Contig ID
                  2534 7.R1040
5'-most EST
                  LIB3028-021-Q1-B1-A6
Seq. No.
                  3541
                  2536 1.R1040
Contig ID
                  rlr700901294.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3241945
BLAST score
                  376
E value
                  6.0e-36
Match length
                  156
% identity
                  48
NCBI Description (AC004625) unknown protein [Arabidopsis thaliana]
Seq. No.
                  2539_1.R1040
Contig ID
5'-most EST
                  uC-qmflminsoy064d02b1
Method
                  BLASTX
NCBI GI
                  q2960364
BLAST score
                  1321
                  1.0e-146
E value
Match length
                  332
                  74
% identity
NCBI Description
                  (AJ224986) cinnamoyl CoA reductase [Populus balsamifera
                  subsp. trichocarpa]
                  3543
Seq. No.
                   2539 2.R1040
Contig ID
5'-most EST
                   jC-qmro02910026a02a1
Method
                  BLASTX
NCBI GI
                  g2960364
BLAST score
                   291
E value
                   3.0e-26
Match length
                  71
% identity
                  75
                  (AJ224986) cinnamoyl CoA reductase [Populus balsamifera
NCBI Description
                   subsp. trichocarpa]
```

Seq. No. 3544

Contig ID 2539 3.R1040

5'-most EST LIB3050-008-Q1-E1-G8

Method BLASTX



```
q2058311
NCBI GI
                  466
BLAST score
                  1.0e-46
E value
                  116
Match length
                  76
% identity
NCBI Description (X79566) cinnamoyl-CoA reductase [Eucalyptus gunnii]
Seq. No.
                  3545
                  2539 4.R1040
Contig ID
                  jsh701068331.h1
5'-most EST
                  BLASTX
Method
                  g2960364
NCBI GI
                  400
BLAST score
                   5.0e-39
E value
                  100
Match length
                   75
% identity
                  (AJ224986) cinnamoyl CoA reductase [Populus balsamifera
NCBI Description
                   subsp. trichocarpa]
                   3546
Seq. No.
                   2539 5.R1040
Contig ID
                   pmv700889373.h1
5'-most EST
                   BLASTX
Method
                   q2960364
NCBI GI
                   451
BLAST score
                   8.0e-45
E value
                   115
Match length
                   75
% identity
                  (AJ224986) cinnamoyl CoA reductase [Populus balsamifera
NCBI Description
                   subsp. trichocarpa]
                   3547
Seq. No.
                   2539 7.R1040
Contig ID
                   jC-qmst02400030a10a1
5'-most EST
                   BLASTX
Method
                   g2058311
NCBI GI
                   190
BLAST score
                   2.0e-29
E value
                   92
Match length
                   72
% identity
NCBI Description (X79566) cinnamoyl-CoA reductase [Eucalyptus gunnii]
Seq. No.
                   3548
                   2540 1.R1040
Contig ID
                   jC-qmf102220068c05a1
 5'-most EST
                   BLASTX
Method
                   q633890
NCBI GI
 BLAST score
                   406
 E value
                   1.0e-39
                   127
 Match length
 % identity
                   (S72926) glucose and ribitol dehydrogenase homolog [Hordeum
 NCBI Description
                   vulgare]
```

3549

Seq. No.

2541 1.R1040 Contig ID

jC-qmf102220091e04a1 5'-most EST



Method BLASTX
NCBI GI g128592
BLAST score 425
E value 8.0e-42
Match length 126
% identity 63

NCBI Description POLLEN-SPECIFIC PROTEIN NTP303 PRECURSOR

>gi_82190_pir__S22495 pollen-specific protein precursor common tobacco >gi_19902_emb_CAA43454_ (X61146) pollen

specific protein [Nicotiana tabacum]

Seq. No. 3550

Contig ID 2544_1.R1040 5'-most EST ujr700646514.h1

Seq. No. 3551

Contig ID 2544_4.R1040 5'-most EST hrw701063009.h1

Seq. No. 3552

Contig ID 2545_1.R1040 5'-most EST sat701014436.h1

Method BLASTX
NCBI GI g2160156
BLAST score 606
E value 1.0e-62
Match length 214
% identity 56

NCBI Description (AC000132) Strong similarity to S. pombe leucyl-tRNA

synthetase (gb Z73100). [Arabidopsis thaliana]

Seq. No. 3553

Contig ID 2547 1.R1040

5'-most EST LIB3107-061-Q1-K1-H1

Seq. No. 3554

Contig ID 2547_2.R1040 5'-most EST pcp700995149.h1

Seq. No. 3555

Contig ID 2548_1.R1040 5'-most EST uC-gmropic074b05b1

Method BLASTX
NCBI GI g2506277
BLAST score 2497
E value 0.0e+00
Match length 543
% identity 93

NCBI Description RUBISCO SUBUNIT BINDING-PROTEIN BETA SUBUNIT PRECURSOR (60

KD CHAPERONIN BETA SUBUNIT) (CPN-60 BETA) >gi 806808

(U21139) chaperonin precursor [Pisum sativum]

Seq. No. 3556

Contig ID 2548 2.R1040

5'-most EST LIB3106-005-Q1-K1-F7

Method BLASTN NCBI GI g2746720



BLAST score 128
E value 2.0e-65
Match length 308
% identity 85
NCBI Description Capsicum annuum histone H4 mRNA, complete cds
Seq. No. 3557
Contig ID 2548_3.R1040
5'-most EST LIB3040-055-Q1-E1-H12

5'-most EST LIB3040-055

Method BLASTN

NCBI GI g2746720

BLAST score 115

E value 9.0e-58

Match length 299

Match length 299 % identity 85

NCBI Description Capsicum annuum histone H4 mRNA, complete cds

Seq. No. 3558

Contig ID 2548_4.R1040

5'-most EST LIB3170-002-Q1-K1-D3

Method BLASTN
NCBI GI g2746720
BLAST score 133
E value 2.0e-68
Match length 309
% identity 86

NCBI Description Capsicum annuum histone H4 mRNA, complete cds

Seq. No. 3559

Contig ID 2548_5.R1040

5'-most EST LIB3040-013-Q1-E1-E12

Method BLASTN
NCBI GI g2351070
BLAST score 106
E value 4.0e-52
Match length 294
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTH12, complete sequence [Arabidopsis thaliana]

Seq. No. 3560

Contig ID 2548_6.R1040

5'-most EST LIB3106-105-Q1-K1-E10

Method BLASTN
NCBI GI g2746720
BLAST score 112
E value 6.0e-56
Match length 308
% identity 84

NCBI Description Capsicum annuum histone H4 mRNA, complete cds

Seq. No. 3561

Contig ID 2548_7.R1040

5'-most EST LIB3106-005-Q1-K2-F7

Method BLASTN
NCBI GI g2746720
BLAST score 117



```
5.0e-59
E value
                  309
Match length
                  84
% identity
                  Capsicum annuum histone H4 mRNA, complete cds
NCBI Description
                  3562
Seq. No.
                  2548 8.R1040
Contig ID
                  zhf700957937.h1
5'-most EST
                  BLASTN
Method
                   g2746720
NCBI GI
                  121
BLAST score
                   2.0e-61
E value
                   309
Match length
                   85
% identity
NCBI Description Capsicum annuum histone H4 mRNA, complete cds
                   3563
Seq. No.
                   2550 1.R1040
Contig ID
                   LIB3106-069-P1-K1-A5
5'-most EST
                   3564
Seq. No.
                   2555 1.R1040
Contig ID
                   LIB3039-039-Q1-E1-H5
5'-most EST
                   BLASTX
Method
                   g3075394
NCBI GI
                   2205
BLAST score
                   0.0e+00
E value
Match length
                   532
                   79
% identity
                   (AC004484) putative beta-ketoacyl-CoA synthase [Arabidopsis
NCBI Description
                   thaliana] >gi_3559809_emb_CAA09311_ (AJ010713) fiddlehead
                   protein [Arabidopsis thaliana]
                   3565
Seq. No.
                   2555 2.R1040
Contig ID
                   leu7\overline{0}1146601.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3075394
BLAST score
                   414
                   1.0e-40
E value
Match length
                   97
                   82
% identity
                   (AC004484) putative beta-ketoacyl-CoA synthase [Arabidopsis
NCBI Description
                   thaliana] >gi_3559809_emb_CAA09311_ (AJ010713) fiddlehead
                   protein [Arabidopsis Thaliana]
                   3566
Seq. No.
Contig ID
                   2556 1.R1040
                   LIB3028-042-Q1-B1-D6
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g3894172
BLAST score 811
E value 2.0e-86
Match length 319
% identity 48

NCBI Description (AC005312) putative cinnamoyl-CoA reductase [Arabidopsis

thaliana]



```
3567
Seq. No.
                   2556 2.R1040
Contig ID
                   LIB3109-037-Q1-K1-E4
5'-most EST
                   BLASTN
Method
                   g2656025
NCBI GI
BLAST score
                   42
                   3.0e-14
E value
                   58
Match length
                   93
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MCD7
                   3568
Seq. No.
                   2556 3.R1040
Contig ID
                   kmv700737795.h1
5'-most EST
                   BLASTX
Method
                   g3894172
NCBI GI
                   171
BLAST score
                   3.0e-12
E value
Match length
                   56
% identity
                   (AC005312) putative cinnamoyl-CoA reductase [Arabidopsis
NCBI Description
                   thaliana]
                   3569
Seq. No.
                   2556 5.R1040
Contig ID
                   uC-gmrominsoy121d12b1
5'-most EST
                   {\tt BLASTN}
Method
                   g2656025
NCBI GI
BLAST score
                   39
                   1.0e-12
E value
                   51
Match length
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MCD7
                   3570
Seq. No.
                   2556 6.R1040
Contig ID
                   LIB3109-018-Q1-K1-H8
5'-most EST
                   BLASTX
Method
                   g3482924
NCBI GI
BLAST score
                   191
E value
                   6.0e - 24
                   117
Match length
% identity
                   (AC003970) Highly similar to cinnamyl alcohol
NCBI Description
                   dehydrogenase, gi_1143445 [Arabidopsis thaliana]
                   3571
Seq. No.
Contig ID
                   2556 7.R1040
                   ssr700560974.h1
5'-most EST
Method
                   BLASTX
```

NCBI GI g3482925
BLAST score 125
E value 2.0e-13
Match length 92



% identity (AC003970) Highly similar to cinnamyl alcohol NCBI Description dehydrogenase, gi 1143445 [Arabidopsis thaliana] Seq. No. 3572 Contig ID 2556 8.R1040 jC-gmle01810036c11d1 5'-most EST Method BLASTX q3894172 NCBI GI 161 BLAST score E value 4.0e-11 67 Match length

% identity (ACO05312) putative cinnamoyl-CoA reductase [Arabidopsis NCBI Description

thaliana]

Seq. No. 3573

Contig ID 2557 1.R1040 epx701108487.hl 5'-most EST

3574 Seq. No.

Contig ID 2558 1.R1040

5'-most EST jC-qmle01810089g08a1

BLASTX Method g4220477 NCBI GI BLAST score 191 E value 3.0e-14 91 Match length % identity

(AC006069) unknown protein [Arabidopsis thaliana] NCBI Description

3575 Seq. No.

2558 2.R1040 Contig ID fde700874755.hl 5'-most EST

BLASTX Method g4220477 NCBI GI BLAST score 180 5.0e-13 E value 89 Match length 51 % identity

(AC006069) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

3576

2560 1.R1040 Contig ID

LIB3028-042-Q1-B1-E1 5'-most EST

Seq. No.

3577

2562 1.R1040 Contig ID 5'-most EST smw700646136.h1

Seq. No.

3578

2566 1.R1040 Contig ID

 $jC-g\overline{m}le01810067e05a1$ 5'-most EST

BLASTX Method q2129578 NCBI GI BLAST score 330 1.0e-30 E value



Match length 133 % identity 54

NCBI Description dTDP-glucose 4-6-dehydratases homolog - Arabidopsis

thaliana >gi_928932_emb_CAA89205_ (Z49239) homolog of dTDP-glucose 4-6-dehydratases [Arabidopsis thaliana] >gi 1585435 prf_2124427B diamide resistance gene

[Arabidopsis thaliana]

Seq. No. 3579

Contig ID 2566_2.R1040

5'-most EST jC-gmro02910030d09a1

Method BLASTX
NCBI GI g3522929
BLAST score 901
E value 1.0e-114
Match length 299

% identity 76

NCBI Description (AC002535) putative dTDP-glucose 4-6-dehydratase

[Arabidopsis thaliana] >gi_3738279 (AC005309) putative dTDP-glucose 4-6-dehydratase [Arabidopsis thaliana]

Seq. No. 3580

Contig ID 2566_3.R1040

5'-most EST LIB3051-019-Q1-E1-G5

Method BLASTX
NCBI GI g2129578
BLAST score 217
E value 2.0e-17
Match length 112
% identity 46

NCBI Description dTDP-glucose 4-6-dehydratases homolog - Arabidopsis

thaliana >gi_928932_emb_CAA89205_ (Z49239) homolog of dTDP-glucose 4-6-dehydratases [Arabidopsis thaliana] >gi_1585435 prf 2124427B diamide resistance gene

[Arabidopsis thaliana]

Seq. No. 3581

Contig ID 2567_1.R1040 5'-most EST xzy700966701.h1

Method BLASTX
NCBI GI g1743354
BLAST score 2238
E value 0.0e+00
Match length 502
% identity 82

NCBI Description (Y09876) aldehyde dehydrogenase (NAD+) [Nicotiana tabacum]

Seq. No. 3582

Contig ID 2567_2.R1040 5'-most EST pmv700890220.h1

Method BLASTX
NCBI GI g1421730
BLAST score 406
E value 2.0e-39
Match length 111
% identity 69

NCBI Description (U43082) RF2 [Zea mays]



```
3583
Seq. No.
                  2567 3.R1040
Contig ID
                  jC-gmro02800029f04d1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g1743354
                  353
BLAST score
                  1.0e-33
E value
                  71
Match length
% identity
                  (Y09876) aldehyde dehydrogenase (NAD+) [Nicotiana tabacum]
NCBI Description
                   3584
Seq. No.
                  2567 5.R1040
Contig ID
                  zsq701126328.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                   g1421730
                   418
BLAST score
                   5.0e-41
E value
                   86
Match length
                   88
% identity
NCBI Description
                  (U43082) RF2 [Zea mays]
                   3585
Seq. No.
                   2569 1.R1040
Contig ID
                   LIB3028-042-Q1-B1-E8
5'-most EST
                   BLASTX
Method
                   g2459445
NCBI GI
                   218
BLAST score
E value
                   1.0e-17
                   96
Match length
                   26
% identity
                   (AC002332) putative ribonucleoprotein [Arabidopsis
NCBI Description
                   thaliana]
                   3586
Seq. No.
Contig ID
                   2570 1.R1040
                   dpv701101023.hl
5'-most EST
Method
                   BLASTX
                   g2959767
NCBI GI
BLAST score
                   1002
                   1.0e-109
E value
Match length
                   261
                   75
% identity
                   (AJ002584) AtMRP4 [Arabidopsis thaliana] >gi 3738292
NCBI Description
                   (AC005309) glutathione-conjugate transporter AtMRP4
                   [Arabidopsis thaliana]
                   3587
Seq. No.
                   2572 1.R1040
Contig ID
                   LIB3028-042-Q1-B1-B7
5'-most EST
```

Seq. No. 3588

Contig ID 2574_1.R1040

5'-most EST LIB3028-042-Q1-B1-C1

Method BLASTX NCBI GI g4490330



```
BLAST score
                  1.0e-128
E value
                  257
Match length
                  79
% identity
                   (AL035656) splicing factor-like protein [Arabidopsis
NCBI Description
                  thaliana]
                   3589
Seq. No.
                  2574 2.R1040
Contig ID
                  crh700856225.h1
5'-most EST
                  BLASTX
Method
                   g4490330
NCBI GI
                   1033
BLAST score
                   1.0e-113
E value
                   220
Match length
                   88
% identity
                   (AL035656) splicing factor-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3590
Seq. No.
                   2577 1.R1040
Contig ID
                   LIB3051-084-Q1-K1-F10
5'-most EST
                   BLASTX
Method
                   g2129918
NCBI GI
BLAST score
                   915
                   1.0e-101
E value
                   405
Match length
                   54
% identity
                  BPF-1 protein - parsley >gi 396197 emb CAA48413 (X68337)
NCBI Description
                   BPF-1 [Petroselinum crispum] >gi 441310 emb CAA44518_
                   (X62653) BPF-1 [Petroselinum crispum]
Seq. No.
                   3591
                   2578 1.R1040
Contig ID
5'-most EST
                   leu701144328.hl
                   BLASTX
Method
NCBI GI
                   g3334667
BLAST score
                   405
E value
                   3.0e-39
Match length
                   214
% identity
                   40
                  (Y10493) putative cytochrome P450 [Glycine max]
NCBI Description
                   3592
Seq. No.
                   2579 1.R1040
Contig ID
                   LIB3170-048-Q1-J1-H6
5'-most EST
                   BLASTX
Method
                   q3355480
NCBI GI
                   159
BLAST score
                   1.0e-10
E value
```

65 Match length 52 % identity

(AC004218) Medicago nodulin N21-like protein [Arabidopsis NCBI Description

thaliana]

3593

Seq. No.

2580_1.R1040 Contig ID

% identity

57



```
5'-most EST
                   jC-gmro02910072d02a1
                  BLASTX
Method
                  g3859659
NCBI GI
                   165
BLAST score
                   2.0e-11
E value
Match length
                   53
% identity
                   51
                   (AL031394) putative potassium transporter AtKT5p (AtKT5)
NCBI Description
                   [Arabidopsis thaliana]
                   3594
Seq. No.
                   2582 1.R1040
Contig ID
                   gsv701050196.h1
5'-most EST
                   BLASTX
Method
                   g4539008
NCBI GI
BLAST score
                   522
                   3.0e-53
E value
                   165
Match length
                   65
% identity
                  (AL049481) lipase-like protein [Arabidopsis thaliana]
NCBI Description
                   3595
Seq. No.
                   2583 1.R1040
Contig ID
                   hyd7\overline{0}0731246.h1
5'-most EST
                   BLASTX
Method
                   q2194119
NCBI GI
BLAST score
                   633
                   5.0e-66
E value
                   237
Match length
% identity
                   51
                  (AC002062) No definition line found [Arabidopsis thaliana]
NCBI Description
                   3596
Seq. No.
                   2584 1.R1040
Contig ID
                   LIB3092-028-Q1-K1-F1
5'-most EST
Method
                   BLASTX
                   q4455313
NCBI GI
                   304
BLAST score
                   3.0e-27
E value
Match length
                   276
% identity
                   29
                   (AL035528) fatty acid elongase-like protein (cer2-like)
NCBI Description
                   [Arabidopsis thaliana]
                   3597
Seq. No.
                   2584 2.R1040
Contig ID
                   LIB3028-042-Q1-B1-D11
5'-most EST
                   3598
Seq. No.
                   2585 1.R1040
Contig ID
5'-most EST
                   LIB3040-033-Q1-E1-F7
Method
                   BLASTX
                   q3367591
NCBI GI
BLAST score
                   393
                   5.0e-38
E value
Match length
                   124
```



```
NCBI Description (AL031135) putative protein [Arabidopsis thaliana]
```

3599 Seq. No.

2588 1.R1040 Contig ID

jC-gmle01810020g09a1 5'-most EST

BLASTX Method g4263711 NCBI GI 647 BLAST score 1.0e-67 E value 170 Match length 71 % identity

(AC006223) putative CCR4-associated transcription factor NCBI Description

[Arabidopsis thaliana]

3600 Seq. No.

2588 2.R1040 Contig ID

5'-most EST uC-gmropic020e10b1

BLASTX Method NCBI GI g4263711 332 BLAST score E value 8.0e-31 82 Match length 73 % identity

(AC006223) putative CCR4-associated transcription factor NCBI Description

[Arabidopsis thaliana]

3601 Seq. No.

2588 3.R1040 Contig ID

LIB3167-002-Q1-K1-A12 5'-most EST

BLASTX Method q4263711 NCBI GI BLAST score 282 E value 3.0e-25 67 Match length

75 % identity

(ACO06223) putative CCR4-associated transcription factor NCBI Description

[Arabidopsis thaliana]

3602 Seq. No.

Contig ID 2588 4.R1040

5'-most EST jC-gmro02800038g09a1

3603

Seq. No.

2589 1.R1040 Contig ID

LIB3138-093-Q1-N1-F1 5'-most EST

Method BLASTX g4539359 NCBI GI BLAST score 272 2.0e-23 E value 168 Match length % identity 36

(AL049525) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 3604

2589 2.R1040 Contig ID leu701152776.h1 5'-most EST



```
3605
Seq. No.
                   2593 1.R1040
Contig ID
5'-most EST
                  uC-gmronoir044b03b1
Method
                  BLASTX
NCBI GI
                  g2129613
BLAST score
                  529
                   1.0e-83
E value
Match length
                   342
% identity
                   48
                  homeotic protein BEL1 - Arabidopsis thaliana >gi_1122533
NCBI Description
                   (U39944) BELL1 [Arabidopsis thaliana]
                   3606
Seq. No.
                   2593 2.R1040
Contig ID
                   fde700873280.h1
5'-most EST
                   3607
Seq. No.
                   2594 1.R1040
Contig ID
5'-most EST
                   fua701038018.h1
Method
                   BLASTX
NCBI GI
                   g3128228
BLAST score
                   514
                   6.0e-52
E value
Match length
                   178
% identity
                   68
                   (AC004077) putative ribosomal protein L18A [Arabidopsis
NCBI Description
                   thaliana] >gi 3337376 (AC004481) putative ribosomal protein
                   L18A [Arabidopsis thaliana]
Seq. No.
                   3608
                   2595 1.R1040
Contig ID
                   LIB3028-042-Q1-B1-A7
5'-most EST
Method
                   BLASTX
                   g4455340
NCBI GI
BLAST score
                   420
                   3.0e-41
E value
                   138
Match length
                   64
% identity
                   (AL035522) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3609
                   2597 1.R1040
Contig ID
5'-most EST
                   LIB3028-042-Q1-B1-A9
Method
                   BLASTX
                   q2832605
NCBI GI
                   175
BLAST score
                   2.0e-12
E value
                   78
Match length
```

% identity
NCBI Description

51 (AL021633) predicted protein [Arabidopsis thaliana]

Seq. No.

3610

Contig ID 5'-most EST

2598_1.R1040 ncj700980293.h1

Method BLASTN NCBI GI g1370199 BLAST score 391



```
0.0e+00
E value
                   611
Match length
                   91
% identity
                  L.japonicus mRNA for small GTP-binding protein, RAC1
NCBI Description
                   3611
Seq. No.
                   2598 2.R1040
Contig ID
                   g4286070
5'-most EST
                   BLASTN
Method
                   g1370199
NCBI GI
                   196
BLAST score
                   1.0e-106
E value
                   276
Match length
% identity
                   93
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAC1
                   3612
Seq. No.
                   2599 1.R1040
Contig ID
                   g4395750
5'-most EST
                   BLASTX
Method
                   g3882309
NCBI GI
BLAST score
                   310
                   3.0e-28
E value
                   186
Match length
% identity
                   34
                   (AB018337) KIAA0794 protein [Homo sapiens]
NCBI Description
                   3613
Seq. No.
                   2609_1.R1040
Contig ID
                   jC-gmst02400027h09a1
5'-most EST
                   BLASTN
Method
NCBI GI
                   q684941
BLAST score
                   352
                   0.0e + 00
E value
                   716
Match length
                   87
% identity
                   Medicago sativa S-adenosyl-L-methionine:trans-caffeoyl-CoA
NCBI Description
                   3-O-methyltransferase (CCOMT) mRNA, complete cds
                    3614
Seq. No.
                    2609 2.R1040
Contig ID
                   jC-gmst02400031b03d1
5'-most EST
Method
                   BLASTN
NCBI GI
                    q857577
BLAST score
                    141
                    3.0e-73
E value
                    313
Match length
                    86
 % identity
                   Populus tremuloides caffeoyl-CoA 3-O-methyltransferase
NCBI Description
                    mRNA, complete cds
                    3615
 Seq. No.
```

Contig ID 2609 3.R1040 $fua7\overline{0}1042628.h1$ 5'-most EST

BLASTN Method g2465011 NCBI GI

96 BLAST score



E value 2.0e-46
Match length 212
% identity 89

NCBI Description Fragaria vesca mRNA for putative caffeoyl-CoA

3-O-methyltransferase, partial

Seq. No. 3616

Contig ID 2609_6.R1040 5'-most EST fde700873805.h1

Method BLASTN
NCBI GI g3550589
BLAST score 36
E value 6.0e-11
Match length 119
% identity 89

NCBI Description Populus trichocarpa CCoAOMT2 gene, exon 1 to exon

Seq. No. 3617

Contig ID 2610 1.R1040

5'-most EST LIB3139-071-P1-N1-G5

Method BLASTX
NCBI GI g4538929
BLAST score 271
E value 9.0e-24
Match length 76
% identity 42

NCBI Description (AL049483) putative nucleic acid binding protein

[Arabidopsis thaliana]

Seq. No. 3618

Contig ID 2613_1.R1040

5'-most EST uC-gmrominsoy280g07b1

Method BLASTX
NCBI GI g3046700
BLAST score 674
E value 1.0e-70
Match length 249
% identity 58

NCBI Description (AJ005261) cytidine deaminase [Arabidopsis thaliana] >gi_3093276_emb_CAA06671_ (AJ005687) cytidine deaminase

[Arabidopsis thaliana] >gi_4191787 (AC005917) putative

cytidine deaminase [Arabidopsis thaliana]

Seq. No. 3619

Contig ID 2614 1.R1040

5'-most EST LIB3028-041-Q1-B1-H3

Method BLASTX

NCBI GI g1076291

BLAST score 204

E value 4.0e-16

Match length 63
% identity 49

NCBI Description amino acid transporter AAT1 - Arabidopsis thaliana

>gi_2911069_emb_CAA17531_ (AL021960) amino acid transport

protein AATI [Arabidopsis thaliana]

Seq. No. 3620



2616 1.R1040 Contig ID $taw7\overline{0}0655504.h1$ 5'-most EST 3621 Seq. No. 2620 1.R1040 Contig ID LIB3028-042-Q1-B1-A1 5'-most EST Method BLASTX NCBI GI g2245068 BLAST score 545 3.0e-55 E value 133 Match length 74 % identity (Z97342) serine protease homolog [Arabidopsis thaliana] NCBI Description 3622 Seq. No. 2620 2.R1040 Contig ID fua701038204.h1 5'-most EST Seq. No. 3623 2620 3.R1040 Contig ID jC-gmst02400045e09a1 5'-most EST 3624 Seq. No. 2620 6.R1040 Contig ID zhf700962164.h15'-most EST BLASTX Method g2245068 NCBI GI 152 BLAST score E value 5.0e-10 39 Match length % identity (Z97342) serine protease homolog [Arabidopsis thaliana] NCBI Description Seq. No. 3625 2621 1.R1040 Contig ID uC-qmropic089d07b1 5'-most EST Method BLASTX NCBI GI q303742 BLAST score 1049 1.0e-114 E value 217 Match length % identity (D12544) GTP-binding protein [Pisum sativum] NCBI Description >qi 738936 prf 2001457D GTP-binding protein [Pisum sativum] Seq. No. 3626 Contig ID 2622 1.R1040 LIB3028-041-Q1-B1-E4 5'-most EST

Seq. No.

3627

2623 1.R1040 Contig ID

5'-most EST LIB3053-004-Q1-N1-E3

BLASTX Method g1076621 NCBI GI 432 BLAST score E value 2.0e-42



```
Match length 140 % identity 56 NCBI Description cyto (X7)
```

cytochrome b5 - common tobacco >gi_296386_emb_CAA50575_
(X71441) cytochrome b5 [Nicotiana tabacum]

Method BLASTX
NCBI GI g2760832
BLAST score 480
E value 3.0e-48
Match length 154
% identity 66

NCBI Description (AC003105) similar to barley ids-4 gene product

[Arabidopsis thaliana]

Seq. No. 3629

Contig ID 2624 2.R1040

5'-most EST uC-gmronoir030b12b1

Method BLASTX
NCBI GI g2760832
BLAST score 373
E value 8.0e-36
Match length 111
% identity 70

NCBI Description (AC003105) similar to barley ids-4 gene product

[Arabidopsis thaliana]

Seq. No. 3630

Contig ID 2625_1.R1040

5'-most EST LIB3039-012-Q1-E1-B7

Method BLASTX
NCBI GI g3643610
BLAST score 972
E value 1.0e-105
Match length 211
% identity 87

NCBI Description (AC005395) putative serine/threonine protein kinase

[Arabidopsis thaliana]

Seq. No. 3631

Contig ID 2627_1.R1040

5'-most EST LIB3093-038-Q1-K1-H11

Method BLASTX
NCBI GI g4263711
BLAST score 1041
E value 1.0e-113
Match length 247
% identity 77

NCBI Description (AC006223) putative CCR4-associated transcription factor

[Arabidopsis thaliana]

Seq. No. 3632

Contig ID 2627 2.R1040 5'-most EST g5058533 Method BLASTX



```
NCBI GI
                   g4106061
                   269
BLAST score
                   2.0e-23
E value
                   83
Match length
% identity
                   61
NCBI Description (AF053318) CCR4-associated factor 1 [Homo sapiens]
Seq. No.
                   3633
                   2627 4.R1040
Contig ID
                   LIB3049-043-Q1-E1-B1
5'-most EST
                   3634
Seq. No.
                   2630 1.R1040
Contig ID
                   pmv700893795.hl
5'-most EST
                   3635
Seq. No.
                   2631 1.R1040
Contig ID
                   uaw7\overline{0}0663361.h1
5'-most EST
                   3636
Seq. No.
                   2632 1.R1040
Contig ID
5'-most EST
                   LIB3170-013-Q1-K1-H5
Seq. No.
                   3637
                   2632 2.R1040
Contig ID
5'-most EST
                   LIB3074-009-Q1-E1-C5
                   3638
Seq. No.
                   2633 1.R1040
Contig ID
                   zhf700953666.h1
5'-most EST
                   3639
Seq. No.
                   2634 1.R1040
Contig ID
                   6HA - \overline{0}2 - Q1 - B1 - A12
5'-most EST
                   3640
Seq. No.
                   2636 1.R1040
Contig ID
                    jC-gmst02400061f01a1
5'-most EST
                   BLASTX
Method
                   q4454452
NCBI GI
BLAST score
                   936
E value
                    1.0e-101
Match length
                   278
% identity
                   (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                    3641
Seq. No.
Contig ID
                    2638 1.R1040
5'-most EST
                   bnu700967569.hl
Method
                    BLASTX
```

NCBI GI q2330797

BLAST score 371 E value 6.0e-35 Match length 205 % identity

NCBI Description (Z98601) zinc finger protein [Schizosaccharomyces pombe]



```
Seq. No.
                  2639 1.R1040
Contig ID
                  ssr700558953.hl
5'-most EST
                  BLASTX
Method
                  g4468986
NCBI GI
                  321
BLAST score
                  2.0e-29
E value
Match length
                  187
                  38
% identity
                   (AL035605) putative protein [Arabidopsis thaliana]
NCBI Description
                  3643
Seq. No.
                  2639 2.R1040
Contig ID
                  ncj700981812.hl
5'-most EST
                  BLASTX
Method
                   g4468986
NCBI GI
                   170
BLAST score
                   4.0e-12
E value
                   93
Match length
                   40
% identity
                  (AL035605) putative protein [Arabidopsis thaliana]
NCBI Description
                   3644
Seq. No.
                   2639 3.R1040
Contig ID
                   awf700839353.hl
5'-most EST
                   BLASTX
Method
                   g4468986
NCBI GI
                   186
BLAST score
                   7.0e-14
E value
                   102
Match length
% identity
                   41
                  (AL035605) putative protein [Arabidopsis thaliana]
NCBI Description
                   3645
Seq. No.
                   2641 1.R1040
Contig ID
                   wvk700684348.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   q602358
BLAST score
                   344
                   0.0e + 00
E value
                   811
Match length
                   89
% identity
NCBI Description P.sativum mRNA for type II chlorophyll a/b binding protein
                   3646
Seq. No.
                   2641 2.R1040
Contig ID
                   LIB3039-028-Q1-E1-G10
5'-most EST
                   BLASTN
Method
NCBI GI
                   q602358
BLAST score
                   260
                   1.0e-144
E value
Match length
                   723
% identity
```

Seq. No. 3647

NCBI Description

Contig ID 2643_1.R1040

P.sativum mRNA for type II chlorophyll a/b binding protein

e, .



```
5'-most EST
                    LIB3138-048-Q1-N1-H4
                    3648
 Seq. No.
                    2643 2.R1040
 Contig ID
                    LIB3092-033-Q1-K1-G10
 5'-most EST
                    BLASTX
 Method
                    g2190540
 NCBI GI
 BLAST score
                    331
                    1.0e-30
 E value
 Match length
                    74
                    80
 % identity
                    (ACO01229) Similar to Arabidopsis TFL1 (gb_U77674).
 NCBI Description
                    [Arabidopsis thaliana]
                    3649
 Seq. No.
                    2644 1.R1040
 Contig ID
                    LIB3039-012-Q1-E1-D3
 5'-most EST
 Seq. No.
                    3650
                    2647 1.R1040
 Contig ID
                    LIB3109-020-Q1-K1-D10
 5'-most EST
 Method
                    BLASTN
                    q169980
 NCBI GI
 BLAST score
                    138
E value
                    3.0e-71
                    447
 Match length
                    87
 % identity
 NCBI Description Soybean heat-shock protein (Gmhsp26-A) gene, complete cds
                    3651
 Seq. No.
                    2651 1.R1040
 Contig ID
 5'-most EST
                    seb700652937.h1
                    BLASTX
 Method
 NCBI GI
                    q2982303
 BLAST score
                    1236
                    1.0e-136
 E value
                    267
 Match length
 % identity
                   (AF051236) hypothetical protein [Picea mariana]
 NCBI Description
                    3652
 Seq. No.
                    2653 1.R1040
 Contig ID
 5'-most EST
                    LIB3093-042-Q1-K1-A6
 Method
                    BLASTX
                    q2501337
 NCBI GI
 BLAST score
                    500
                    1.0e-119
 E value
                    414
 Match length
  % identity
                    52
                    AMINE OXIDASE PRECURSOR [COPPER-CONTAINING]
 NCBI Description
                    >gi_2129875_pir__C44239 amine oxidase (copper-containing)
                    (EC 1.4.3.6) precursor - garden pea >gi_685198 (L39931)
                    copper amine oxidase [Pisum sativum]
                    3653
```

Seq. No.

2656 1.R1040 Contig ID

uC-gmrominsoy026h04b1 5'-most EST



```
BLASTX
Method
                   g2275196
NCBI GI
                   1337
BLAST score
                   1.0e-148
E value
                   328
Match length
                   77
% identity
                   (AC002337) water stress-induced protein, WSI76 isolog
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   3654
                   2658 1.R1040
Contig ID
                   LIB3049-031-Q1-E1-C9
5'-most EST
                   BLASTX
Method
                   q4559346
NCBI GI
                   400
BLAST score
                   1.0e-38
E value
                   176
Match length
                   45
% identity
                   (AC006585) early nodulin 16 [Arabidopsis thaliana]
NCBI Description
                   3655
Seq. No.
                   2660 1.R1040
Contig ID
                   uxk7\overline{0}0671869.h1
5'-most EST
                   BLASTX
Method
                   g2088652
NCBI GI
BLAST score
                   671
                   1.0e-70
E value
                   166
Match length
% identity
                   74
                   (AF002109) 26S proteasome regulatory subunit S12 isolog
NCBI Description
                    [Arabidopsis thaliana] >gi_2351376 (U54561) translation
                   initiation factor eIF2 p47 subunit homolog [Arabidopsis
                   thaliana]
Seq. No.
                   3656
                   2661 1.R1040
Contig ID
                   g4292101
5'-most EST
                   BLASTX
Method
                   g4262149
NCBI GI
BLAST score
                    607
                    5.0e-63
E value
                   151
Match length
                    73
 % identity
                    (AC005275) putative xyloglucan endotransglycosylase
NCBI Description
                    [Arabidopsis thaliana]
                    3657
 Seq. No.
                    2661 2.R1040
 Contig ID
 5'-most EST
                    fC-gmle700554113a3
                    BLASTX
Method
                    q4262149
 NCBI GI
 BLAST score
                    622
```

8.0e-65 E value

Match length 136 % identity

(AC005275) putative xyloglucan endotransglycosylase NCBI Description

[Arabidopsis thaliana]



```
3658
Seq. No.
                   2663 1.R1040
Contig ID
                   LIB3028-040-Q1-B1-H3
5'-most EST
                   3659
Seq. No.
                   2664 1.R1040
Contig ID
                   fde7\overline{0}0874478.h1
5'-most EST
                   BLASTX
Method
                   q1922240
NCBI GI
                   143
BLAST score
                   1.0e-08
E value
                   28
Match length
% identity
                   82
NCBI Description (Y10083) hypothetical protein [Arabidopsis thaliana]
                   3660
Seq. No.
                   2669 1.R1040
Contig ID
                   LIB3049-034-Q1-E1-F6
5'-most EST
                   BLASTX
Method
                   q4454263
NCBI GI
                   152
BLAST score
                   2.0e-09
E value
                   94
Match length
% identity
                   38
NCBI Description (Y18046) FGFR1 oncogene partner (FOP) [Homo sapiens]
                   3661
Seq. No.
                   2669 2.R1040
Contig ID
                   vzy700755346.h1
5'-most EST
                   {\tt BLASTX}
Method
                   q4454263
NCBI GI
                   153
BLAST score
E value
                   5.0e-10
                   94
Match length
                   38
% identity
                   (Y18046) FGFR1 oncogene partner (FOP) [Homo sapiens]
NCBI Description
Seq. No.
                   3662
                   2669 3.R1040
Contig ID
                   eep700864394.h1
5'-most EST
                    3663
Seq. No.
Contig ID
                    2672 1.R1040
                   uC-qmflminsoy029c04b1
5'-most EST
                   BLASTX
Method
                    a2578444
NCBI GI
                    275
BLAST score
E value
                    1.0e-23
                    89
Match length
 % identity
                   (X67427) ptxA [Pisum sativum]
NCBI Description
 Seq. No.
                    3664
                    2672 2.R1040
 Contig ID
                    jC-qmst02400061d09d1
 5'-most EST
```

```
3665
Seq. No.
                  2672 3.R1040
Contig ID
                  LIB3028-041-Q1-B1-A4
5'-most EST
Method
                  BLASTN
NCBI GI
                  g2578443
BLAST score
                  45
E value
                  6.0e-16
Match length
                  257
% identity
                   76
NCBI Description Pisum sativum ptxA gene
                   3666
Seq. No.
                   2675 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy298a12b1
Method
                   BLASTX
NCBI GI
                   q3319340
BLAST score
                   716
E value
                   2.0e-75
Match length
                   208
% identity
                   67
                   (AF077407) contains similarity to E. coli cation transport
NCBI Description
                   protein ChaC (GB:D90756) [Arabidopsis thaliana]
```

 Seq. No.
 3667

 Contig ID
 2675_2.R1040

 5'-most EST
 leu701148029.h1

 Method
 BLASTX

Method BLASTX
NCBI GI g2827524
BLAST score 615
E value 3.0e-68
Match length 212
% identity 68

NCBI Description (AL021633) predicted protein [Arabidopsis thaliana]

 Seq. No.
 3668

 Contig ID
 2678_1.R1040

 5'-most EST
 LIB3106-036-Q1-K1-D4

 Method
 BLASTX

 NCBI GI
 g3874214

 BLAST score
 647

BLAST score 647
E value 2.0e-67
Match length 361
% identity 42
NCRI Description (783217)

NCBI Description (Z83217) Similarity to Yeast E1-E2 ATPase YEL031W (SW:YED1 YEAST); cDNA EST EMBL:D27574 comes from this gene;

CDNA EST EMBL:D33757 comes from this gene; cDNA EST EMBL:D34256 comes from this gene; cDNA EST EMBL:D37288

comes from

Seq. No. 3669

Contig ID 2679_1.R1040 5'-most EST pcp700993216.h1

Method BLASTX
NCBI GI g3717946
BLAST score 257
E value 7.0e-22
Match length 110

```
% identity
                  (AJ005901) vag1 [Arabidopsis thaliana]
NCBI Description
                  3670
Seq. No.
                  2679 2.R1040
Contig ID
                  taw700654871.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3492806
BLAST score
                  678
E value
                   6.0e-71
                   289
Match length
                   45
% identity
                   (AJ225045) adventitious rooting related oxygenase [Malus
NCBI Description
                   domestica]
                   3671
```

2681 1.R1040 Contig ID LIB3050-004-Q1-E1-A9 5'-most EST Method BLASTX g2961372 NCBI GI 1239 BLAST score E value 1.0e-137

238 Match length 95 % identity

Seq. No.

(AL022141) putative ribosomal protein L8 [Arabidopsis NCBI Description thaliana] >gi 3036817 emb CAA18507 (AL022373) ribosomal

protein L2 [Arabidopsis thaliana]

3672 Seq. No. 2681 2.R1040 Contig ID

LIB3051-063-Q1-K1-G12 5'-most EST

BLASTX Method g2961372 NCBI GI BLAST score 861 8.0e-93 E value 173 Match length 91 % identity

(AL022141) putative ribosomal protein L8 [Arabidopsis NCBI Description thaliana] >qi 3036817 emb CAA18507 (AL022373) ribosomal

protein L2 [Arabidopsis thaliana]

3673 Seq. No.

2681 3.R1040 Contig ID

5'-most EST LIB3170-034-Q1-K1-B9

BLASTN Method q2961370 NCBI GI BLAST score 121 2.0e-61 E value Match length 305 85 % identity

Arabidopsis thaliana DNA chromosome 4, BAC clone F23E13 NCBI Description

(ESSAII project)

Seq. No. 3674

2681 4.R1040 Contig ID

LIB3139-080-P1-N1-B12 5'-most EST

BLASTX Method



q2961372 NCBI GI 416 BLAST score 7.0e-41 E value Match length 132 % identity 71 (AL022141) putative ribosomal protein L8 [Arabidopsis NCBI Description thaliana] >gi 3036817 emb CAA18507 (AL022373) ribosomal protein L2 [Arabidopsis thaliana] 3675 Seq. No. 2681 5.R1040 Contig ID LIB3170-034-Q1-K1-C1 5'-most EST BLASTN Method q4406805 NCBI GI BLAST score 87 5.0e-41 E value 263 Match length 83 % identity Arabidopsis thaliana chromosome II BAC T27K22 genomic NCBI Description sequence, complete sequence [Arabidopsis thaliana] 3676 Seq. No. Contig ID 2681_6.R1040 5'-most EST jC-gmro02910025d01a1 BLASTN Method NCBI GI q4406805 81 BLAST score 1.0e-37 E value Match length 141 89 % identity Arabidopsis thaliana chromosome II BAC T27K22 genomic NCBI Description sequence, complete sequence [Arabidopsis thaliana] Seq. No. 3677 2681 7.R1040 Contig ID $jsh7\overline{0}1067395.h1$ 5'-most EST BLASTX Method g2961372 NCBI GI BLAST score 350 2.0e-33 E value 69 Match length % identity (AL022141) putative ribosomal protein L8 [Arabidopsis NCBI Description thaliana] >gi_3036817_emb_CAA18507_ (AL022373) ribosomal protein L2 [Arabidopsis thaliana] 3678 Seq. No. 2681 8.R1040 Contig ID $epx7\overline{0}1109002.h1$ 5'-most EST BLASTX Method q2961372 NCBI GI BLAST score 359 2.0e-34 E value

Match length 74 % identity

(AL022141) putative ribosomal protein L8 [Arabidopsis NCBI Description thaliana] >gi_3036817_emb_CAA18507_ (AL022373) ribosomal



protein L2 [Arabidopsis thaliana]

Seq. No. 3679

Contig ID 2683_1.R1040

5'-most EST jC-gmfl02220091c10d1

Seq. No. 3680

Contig ID 2686_1.R1040 5'-most EST txt700736173.h1

Method BLASTX
NCBI GI g3182981
BLAST score 2770
E value 0.0e+00
Match length 567
% identity 89

NCBI Description CELL ELONGATION PROTEIN DIMINUTO >gi_1695692_dbj_BAA13096_

(D86494) diminuto [Pisum sativum]

Seq. No. 3681

Contig ID 2690_1.R1040 5'-most EST crh700855489.h1

Method BLASTX
NCBI GI g3510538
BLAST score 403
E value 4.0e-39
Match length 125
% identity 60

NCBI Description (U93167) expansin [Prunus armeniaca]

Seq. No. 3682

Contig ID 2693_1.R1040

5'-most EST LIB3028-040-Q1-B1-F5

Seq. No. 3683

Contig ID 2693_2.R1040

5'-most EST LIB3055-011-Q1-N1-D11

Seq. No. 3684

Contig ID 2694_1.R1040

5'-most EST uC-gmropic011g06b1

Method BLASTX
NCBI GI g2351378
BLAST score 367
E value 1.0e-34
Match length 173
% identity 53

NCBI Description (U54558) translation initiation factor eIF3 p66 subunit

[Homo sapiens] >gi_4200328_emb_CAA18440_ (AL022313)

EIF3-P66 [Homo sapiens]

>gi 4503523 ref_NP_003744.1_pEIF3S7_ UNKNOWN

Seq. No. 3685

Contig ID 2695_1.R1040 5'-most EST vzy700752220.h1

Method BLASTX
NCBI GI g3402282
BLAST score 184



E value 2.0e-28
Match length 112
% identity 54

NCBI Description (AJ000997) proline-rich protein [Solanum tuberosum]

Seq. No. 3686

Contig ID 2699_1.R1040 5'-most EST bth700845695.h1

Method BLASTX
NCBI GI g4512651
BLAST score 1083
E value 1.0e-118
Match length 410
% identity 49

NCBI Description (AC007048) putative tyrosine transaminase [Arabidopsis

thaliana]

Seq. No. 3687

Contig ID 2699_2.R1040 5'-most EST fC-gmf1700908904a1

Seq. No. 3688

Contig ID 2699_3.R1040

5'-most EST jC-gmf102220063e05d1

Method BLASTX
NCBI GI g4512651
BLAST score 263
E value 1.0e-22
Match length 121
% identity 39

NCBI Description (AC007048) putative tyrosine transaminase [Arabidopsis

thaliana]

Seq. No. 3689

Contig ID 2699_4.R1040

5'-most EST jC-gmf102220148b08a1

Method BLASTX
NCBI GI g4512651
BLAST score 203
E value 7.0e-16
Match length 111
% identity 35

NCBI Description (AC007048) putative tyrosine transaminase [Arabidopsis

thaliana]

Seq. No. 3690

Contig ID 2699_5.R1040 5'-most EST pcp700990102.h1

Seq. No. 3691

Contig ID 2700_1.R1040

5'-most EST LIB3028-040-Q1-B1-E5

Seq. No. 3692

Contig ID 2701_2.R1040

5'-most EST LIB3028-040-Q1-B1-E6

Method BLASTX



NCBI GI g4558548 BLAST score 188 E value 6.0e-14 Match length 55 % identity 64

NCBI Description (AC007138) putative RNaseP-associated protein [Arabidopsis

thaliana]

Seq. No. 3693

Contig ID 2701_3.R1040 5'-most EST ejt700606219.h1

Method BLASTX
NCBI GI g4558548
BLAST score 314
E value 2.0e-34
Match length 152
% identity 48

NCBI Description (AC007138) putative RNaseP-associated protein [Arabidopsis

thaliana]

Seq. No. 3694

Contig ID 2703 1.R1040

5'-most EST LIB3107-025-Q1-K1-A8

Seq. No. 3695

Contig ID 2704 1.R1040

5'-most EST uC-gmropic062g06b1

Method BLASTX
NCBI GI g4103243
BLAST score 460
E value 1.0e-48
Match length 185
% identity 59

NCBI Description (AF022368) BIPOSTO [Arabidopsis thaliana]

Seq. No. 3696

Contig ID 2706_1.R1040 5'-most EST zhf700962987.h1

Method BLASTX
NCBI GI g4056489
BLAST score 429
E value 5.0e-42
Match length 222
% identity 42

NCBI Description (AC005896) putative white protein [Arabidopsis thaliana]

Seq. No. 3697

Contig ID 2708_1.R1040 5'-most EST zpv700763310.h1

Method BLASTX
NCBI GI g3873710
BLAST score 353
E value 5.0e-33
Match length 113
% identity 56

NCBI Description (Z73102) predicted using Genefinder; similar to Zinc finger, C2H2 type; cDNA EST EMBL:M89161 comes from this



gene; cDNA EST EMBL:T01394 comes from this gene; cDNA EST EMBL:T02192 comes from this gene; cDNA EST EMBL:D71409 comes

Seq. No. 3698

Contig ID 2708_2.R1040

5'-most EST LIB3093-031-Q1-K1-B12

Method BLASTN
NCBI GI g4191760
BLAST score 54
E value 1.0e-21

Match length 134 % identity 85

NCBI Description Genomic sequence for Arabidopsis thaliana BAC F17F8,

complete sequence [Arabidopsis thaliana]

Seq. No. 3699

Contig ID 2709_1.R1040
5'-most EST g5057680

Method BLASTX

NCBI GI g3033380

BLAST score 1159

E value 1.0e-127

Match length 285 % identity 76

NCBI Description (AC004238) putative coatomer epsilon subunit [Arabidopsis

thaliana]

Seq. No. 3700

Contig ID 2709_2.R1040 5'-most EST gsv701045190.h1

Method BLASTX
NCBI GI g3033380
BLAST score 286
E value 1.0e-25
Match length 83

% identity 67
NCBI Description (AC004238) putative coatomer epsilon subunit [Arabidopsis

thaliana]

Seq. No. 3701

Contig ID 2709_3.R1040 5'-most EST sat701008044.h1

Method BLASTX
NCBI GI g3033380
BLAST score 203
E value 5.0e-16
Match length 45
% identity 82

NCBI Description (AC004238) putative coatomer epsilon subunit [Arabidopsis

thaliana]

3702

Seq. No.

Contig ID 2709_4.R1040 5'-most EST zhf700961443.h1

Seq. No. 3703

NCBI GI

E value

BLAST score



```
2710 1.R1040
Contig ID
                   wvk7\overline{0}0681157.h2
5'-most EST
                   3704
Seq. No.
                   2711 1.R1040
Contig ID
                   uC-gmflminsoy054f05b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4101589
                   629
BLAST score
E value
                   2.0e-65
                   238
Match length
% identity
                   52
NCBI Description (AF005050) aspartyl aminopeptidase [Homo sapiens]
                   3705
Seq. No.
                   2712 1.R1040
Contig ID
                   LIB3093-044-Q1-K1-D4
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2462746
BLAST score
                   1864
                   0.0e + 00
E value
Match length
                   423
% identity
                   84
                   (AC002292) Similar to ATP-citrate-lyase [Arabidopsis
NCBI Description
                   thaliana]
                   3706
Seq. No.
                   2712 2.R1040
Contig ID
                   zhf700959649.h1
5'-most EST
                   BLASTX
Method
                   g2462746
NCBI GI
                   355
BLAST score
                   9.0e-34
E value
Match length
                   88
                   77
% identity
                   (AC002292) Similar to ATP-citrate-lyase [Arabidopsis
NCBI Description
                   thaliana]
                   3707
Seq. No.
                   2712 4.R1040
Contig ID
                   LIB3106-025-Q1-K1-G7
5'-most EST
                   BLASTX
Method
                   q2462746
NCBI GI
BLAST score
                   150
                   6.0e-10
E value
Match length
                   37
% identity
                   (ACO02292) Similar to ATP-citrate-lyase [Arabidopsis
NCBI Description
                   thaliana]
                   3708
Seq. No.
                   2712 6.R1040
Contig ID
5'-most EST
                   wrq700791527.hl
Method
                   BLASTX
```

744

q2462746

5.0e-22

253



Match length 60 % identity 78

NCBI Description (AC002292) Similar to ATP-citrate-lyase [Arabidopsis

thaliana]

Seq. No. 3709

Contig ID 2712_9.R1040 5'-most EST hyd700730039.h1

Method BLASTX
NCBI GI g2462746
BLAST score 167
E value 9.0e-12
Match length 43
% identity 77

NCBI Description (AC002292) Similar to ATP-citrate-lyase [Arabidopsis

thaliana]

Seq. No. 3710

Contig ID 2713_1.R1040 5'-most EST zhf700965295.h1

Method BLASTX
NCBI GI g4006878
BLAST score 1192
E value 1.0e-131
Match length 312
% identity 69

NCBI Description (Z99707) MAP3K-like protein kinase [Arabidopsis thaliana]

Seq. No. 3711

Contig ID 2716_1.R1040 5'-most EST leu701152981.h1

Method BLASTX
NCBI GI g4191774
BLAST score 487
E value 8.0e-49
Match length 168
% identity 60

NCBI Description (AC005917) putative beta-1,3-endoglucanase [Arabidopsis

thaliana]

Seq. No. 3712

Contig ID 2718_1.R1040

5'-most EST LIB3028-001-Q1-B1-B3

Method BLASTX
NCBI GI g4580389
BLAST score 801
E value 1.0e-85
Match length 211
% identity 70

NCBI Description (AC007171) unknown protein [Arabidopsis thaliana]

Seq. No. 3713

Contig ID 2719_1.R1040

5'-most EST jC-gmle01810005c04a1

Method BLASTX NCBI GI g1076685 BLAST score 353



8.0e-33 E value 149 Match length 39 % identity SPF1 protein - sweet potato >gi_484261_dbj_BAA06278 NCBI Description (D30038) SPF1 protein [Ipomoea batatas] 3714 Seq. No. Contig ID 2719 2.R1040 LIB3139-045-P1-N1-E11 5'-most EST BLASTX Method q1076685 NCBI GI 1414 BLAST score 1.0e-157 E value Match length 524 57 % identity SPF1 protein - sweet potato >gi_484261_dbj_BAA06278_ NCBI Description (D30038) SPF1 protein [Ipomoea batatas] Seq. No. 3715 2719 3.R1040 Contig ID jC-gmro02910075a07a1 5'-most EST BLASTX Method q1076685 NCBI GI 468 BLAST score E value 1.0e-46 232 Match length 35 % identity SPF1 protein - sweet potato >gi_484261_dbj_BAA06278_ NCBI Description (D30038) SPF1 protein [Ipomoea batatas] 3716 Seq. No. 2723 1.R1040 Contig ID gsv701045628.h1 5'-most EST BLASTX Method q1769887 NCBI GI 1541 BLAST score 1.0e-172 E value 408 Match length % identity 72 (X95736) amino acid permease 6 [Arabidopsis thaliana] NCBI Description 3717 Seq. No. 2723 2.R1040 Contig ID 5'-most EST LIB3106-018-Q1-K1-D1 BLASTX Method q1769887 NCBI GI BLAST score 443 9.0e-44E value Match length 143 55 % identity (X95736) amino acid permease 6 [Arabidopsis thaliana] NCBI Description Seq. No. 3718

2726 1.R1040 Contig ID sat701014367.hl 5'-most EST

BLASTX Method q4455159 NCBI GI



```
214
BLAST score
                  6.0e-17
E value
                  61
Match length
                  62
% identity
                  (AL021687) putative protein [Arabidopsis thaliana]
NCBI Description
                   3719
Seq. No.
                   2728 1.R1040
Contig ID
                  LIB3049-015-Q1-E1-A2
5'-most EST
                  BLASTX
Method
NCBI GI
                   q2827710
                   214
BLAST score
                   3.0e-17
E value
                   99
Match length
                   45
% identity
                   (AL021684) lysosomal Pro-X carboxypeptidase - like protein
NCBI Description
                   [Arabidopsis thaliana]
                   3720
Seq. No.
                   2732 1.R1040
Contig ID
                   LIB3028-035-Q1-B1-F7
5'-most EST
Seq. No.
                   3721
                   2733 1.R1040
Contig ID
                   LIB3109-014-Q1-K1-D2
5'-most EST
Method
                   BLASTX
                   q3044214
NCBI GI
BLAST score
                   1125
                   1.0e-123
E value
Match length
                   271
% identity
                   (AF057044) acyl-CoA oxidase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3722
                   2733 2.R1040
Contig ID
                   uC-gmrominsoy189c09b1
5'-most EST
                   BLASTX
Method
                   g2245030
NCBI GI
BLAST score
                   260
                   4.0e-24
E value
                   141
Match length
                   68
% identity
                   (Z97341) apetala2 domain TINY homolog [Arabidopsis
NCBI Description
                   thaliana]
                    3723
 Seq. No.
                    2736 1.R1040
 Contig ID
                   pcp700995727.h1
 5'-most EST
                   BLASTX
Method
                   q3335359
 NCBI GI
                    1989
 BLAST score
                    0.0e+00
 E value
                    489
 Match length
 % identity
                    77
                   (AC003028) unknown protein [Arabidopsis thaliana]
 NCBI Description
```

Seq. No. 3724



Contig ID 2736_3.R1040 5'-most EST ujr700646525.h1

Method BLASTX
NCBI GI g3335359
BLAST score 238
E value 3.0e-20
Match length 62
% identity 73

NCBI Description (AC003028) unknown protein [Arabidopsis thaliana]

Seq. No. 3725

Contig ID 2741_1.R1040 5'-most EST uC-gmropic037d01b1

Method BLASTX
NCBI GI 94006886
BLAST score 209
E value 1.0e-16
Match length 107
% identity 48

NCBI Description (Z99708) putative protein [Arabidopsis thaliana]

Seq. No. 3726

Contig ID 2741_2.R1040

5'-most EST uC-gmropic075f05b1

Method BLASTX
NCBI GI g4006886
BLAST score 169
E value 5.0e-12
Match length 72
% identity 51

NCBI Description (Z99708) putative protein [Arabidopsis thaliana]

Seq. No. 3727

Contig ID 2742_1.R1040

5'-most EST LIB3170-015-Q1-K1-C1

Method BLASTX
NCBI GI g3123230
BLAST score 1359
E value 1.0e-150
Match length 657
% identity 42

NCBI Description EUKARYOTIC TRANSLATION INITIATION FACTOR 3 BETA SUBUNIT (EIF-3 BETA) (EIF3 P116) (EIF3 P110) >gi 2558668 (U78525)

eukaryotic translation initiation factor [Homo sapiens]

Seq. No. 3728

Contig ID 2742_2.R1040 5'-most EST gsv701045433.h1

Seq. No. 3729

Contig ID 2743_1.R1040

5'-most EST LIB3028-039-Q1-B2-F8

Seq. No. 3730

Contig ID 2749 1.R1040

5'-most EST LIB3039-028-Q1-E1-B11

Method BLASTX

q2262116 NCBI GI BLAST score 363 2.0e-34 E value Match length 188 37 % identity (AC002343) cellulose synthase isolog [Arabidopsis thaliana] NCBI Description Seq. No. 2749 2.R1040 Contig ID uC-gmflminsoy061b01b1 5'-most EST Method BLASTX NCBI GI g2262116 BLAST score 162 4.0e-11 E value 97 Match length 32 % identity NCBI Description (AC002343) cellulose synthase isolog [Arabidopsis thaliana] 3732 Seq. No. 2752 1.R1040 Contig ID awf700840376.hl 5'-most EST BLASTX Method NCBI GI g2529686 BLAST score 283 E value 3.0e - 33167 Match length % identity 47 (AC002535) putative G-beta-repeat containing protein, 5' NCBI Description partial [Arabidopsis thaliana]

 Seq. No.
 3733

 Contig ID
 2753_1.R1040

 5'-most EST
 smw700646173.h1

 Method
 BLASTX

 MCBL GI
 373406

NCBI GI 9479406
BLAST score 1254
E value 1.0e-138
Match length 253
% identity 92

NCBI Description chlorophyll a/b-binding protein - garden pea

 $>gi_20671_emb_CAA49149_ (X69215)$ chlorophyll a/b-binding

protein [Pisum sativum]

Seq. No. 3734

Contig ID 2753_2.R1040 5'-most EST hyd700729030.h1

Seq. No. 3735

Contig ID 2754 1.R1040

5'-most EST LIB3028-039-Q1-B1-A7

Method BLASTX
NCBI GI g3850818
BLAST score 781
E value 3.0e-83
Match length 190
% identity 75

NCBI Description (Y18349) U2 snRNP auxiliary factor, small subunit [Oryza



```
sativa]
                  3736
Seq. No.
                  2754 2.R1040
Contig ID
                  LIB3138-015-Q1-N2-H12
5'-most EST
                  BLASTX
Method
                  q3850819
NCBI GI
BLAST score
                  790
E value
                   2.0e-84
Match length
                  184
                   77
% identity
                   (Y18349) U2 snRNP auxiliary factor, small subunit [Oryza
NCBI Description
                   sativa]
                   3737
Seq. No.
                   2754 3.R1040
Contig ID
                   seb700651223.h1
5'-most EST
                   3738
Seq. No.
                   2755 1.R1040
Contig ID
                   leu701149305.h1
5'-most EST
Method
                   BLASTX
                   g3236248
NCBI GI
BLAST score
                   1064
E value
                   1.0e-116
                   248
Match length
                   82
% identity
                  (AC004684) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3739
                   2755 2.R1040
Contig ID
                   zhf700953477.hl
5'-most EST
                   BLASTX
Method
                   g3236248
NCBI GI
BLAST score
                   324
                   5.0e-30
E value
                   95
Match length
% identity
                  (AC004684) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3740
Seq. No.
                   2756 1.R1040
Contig ID
5'-most EST
                   LIB3106-090-Q1-K1-H6
Method
                   BLASTX
                   q400578
NCBI GI
                   269
BLAST score
                   2.0e-23
E value
                   109
Match length
% identity
                   46
                   NADH-UBIQUINONE OXIDOREDUCTASE 18 KD SUBUNIT PRECURSOR
NCBI Description
                   (COMPLEX I-18 KD) (CI-18 KD) (COMPLEX I-AQDQ) (CI-AQDQ)
                   >gi_346531_pir__S28240 NADH dehydrogenase (ubiquinone) (EC
                   1.6.5.3) chain CI-18(IP) - bovine >gi_226_emb_CAA44900_
                   (X63215) NADH dehydrogenase [Bos taurus]
```

Seq. No. 3741

Contig ID 2760 1.R1040



```
uC-gmrominsoy265a09b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2281115
BLAST score
                  2848
                  0.0e+00
E value
Match length
                  740
                  75
% identity
                  (AC002330) putative cullin-like 1 protein [Arabidopsis
NCBI Description
                  thaliana]
                  3742
Seq. No.
                  2760 2.R1040
Contig ID
                  g4290426
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2281115
BLAST score
                  407
E value
                  2.0e-39
Match length
                  111
% identity
                  73
                  (AC002330) putative cullin-like 1 protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                   3743
                  2760 3.R1040
Contig ID
5'-most EST
                  uC-gmropic047g08b1
                  BLASTX
Method
NCBI GI
                  g2281115
                  325
BLAST score
                   3.0e-30
E value
Match length
                  68
% identity
                  87
                  (AC002330) putative cullin-like 1 protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                   3744
Contig ID
                   2760 4.R1040
5'-most EST
                  LIB3138-095-Q1-N1-B3
Seq. No.
                   3745
Contig ID
                   2761 1.R1040
5'-most EST
                   LIB3039-040-Q1-E1-E10
Method
                  BLASTX
NCBI GI
                   g473949
BLAST score
                   237
                   6.0e-30
E value
Match length
                   172
% identity
                   44
                  (D29958) KIAA0116 [Homo sapiens]
NCBI Description
                   3746
Seq. No.
                   2761 2.R1040
Contig ID
```

Contig ID 2761_2.R1040 5'-most EST eep700868836.h1

Seq. No. 3747

Contig ID 2765 1.R1040

5'-most EST jC-gmst02400077a08d1

Method BLASTN

751



```
g1706955
NCBI GI
BLAST score
                  81
                  1.0e-37
E value
                  157
Match length
                  88
% identity
                  Gossypium hirsutum cellulose synthase (celA1) mRNA,
NCBI Description
                  complete cds
                  3748
Seq. No.
                  2768 1.R1040
Contig ID
                  LIB3028-039-Q1-B1-F10
5'-most EST
                  BLASTX
Method
                  q2500036
NCBI GI
BLAST score
                  141
                   1.0e-08
E value
                   31
Match length
                   74
% identity
                  DIHYDROOROTASE PRECURSOR (DHOASE) >gi_2121273 (AF000146)
NCBI Description
                   dihydroorotase [Arabidopsis thaliana]
                   >qi 3292818 emb CAA19808_ (AL031018) dihydroorotase
                   [Arabidopsis thaliana]
                   3749
Seq. No.
                   2769 1.R1040
Contig ID
                   jC-gmf102220053a03a1
5'-most EST
                   BLASTX
Method
                   g2388689
NCBI GI
                   359
BLAST score
                   1.0e-101
E value
                   319
Match length
                   66
% identity
                  (AF016633) GH1 protein [Glycine max]
NCBI Description
                   3750
Seq. No.
                   2769 2.R1040
Contig ID
                   zhf700958581.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2388689
                   974
BLAST score
E value
                   1.0e-105
                   324
Match length
% identity
                  (AF016633) GH1 protein [Glycine max]
NCBI Description
                   3751
Seq. No.
                   2769 4.R1040
Contiq ID
                   jC-qmf102220065e10a1
5'-most EST
                   BLASTX
Method
                   g2388689
NCBI GI
BLAST score
                   154
                   4.0e-10
E value
```

Match length 77 % identity 47

(AF016633) GH1 protein [Glycine max] NCBI Description

3752 Seq. No.

2772_1.R1040 Contig ID



```
5'-most EST
                  zhf700960218.h1
                  3753
Seq. No.
                  2774 1.R1040
Contig ID
                  LIB3051-015-Q1-E1-C7
5'-most EST
                  BLASTN
Method
NCBI GI
                  q3334755
BLAST score
                  131
                  6.0e-67
E value
Match length
                  943
                  84
% identity
                  Medicago sativa mRNA for putative arginine/serine-rich
NCBI Description
                  splicing factor
                  3754
Seq. No.
                  2774 2.R1040
Contig ID
5'-most EST
                  LIB3106-093-Q1-K1-E8
                  BLASTN
Method
NCBI GI
                  q3334755
BLAST score
                  242
                  1.0e-133
E value
                  573
Match length
                  86
% identity
                  Medicago sativa mRNA for putative arginine/serine-rich
NCBI Description
                  splicing factor
                   3755
Seq. No.
                  2776 1.R1040
Contig ID
                  LIB3093-050-Q1-K1-D10
5'-most EST
                  BLASTX
Method
NCBI GI
                   g4512624
BLAST score
                  1923
                   0.0e + 00
E value
                   500
Match length
                   73
% identity
                   (AC004793) Strong similarity to gi 3033401 F19I3.29
NCBI Description
                   putative potassium transporter from Arabidopsis thaliana
                   BAC gb_AC004238
                   3756
Seq. No.
                   2779 1.R1040
Contig ID
5'-most EST
                   zhf700956246.h1
Method
                   BLASTX
NCBI GI
                   g2213610
BLAST score
                   278
                   2.0e-24
E value
Match length
                   158
                   23
% identity
                  (AC000103) F21J9.4 [Arabidopsis thaliana]
NCBI Description
                   3757
Seq. No.
                   2786 1.R1040
Contig ID
```

5'-most EST LIB3073-017-Q1-K1-H10

Method BLASTX q3355483 NCBI GI BLAST score 273 6.0e-24 E value



Match length 59 % identity 71 NCBI Description (AC

(AC004218) gibberellin-regulated protein (GASA5)-like [Arabidopsis thaliana]

Seq. No. 3758

Contig ID 2794_1.R1040

5'-most EST LIB3028-039-Q1-B2-A2

Seq. No. 3759

Contig ID 2794 2.R1040

5'-most EST LIB3028-039-Q1-B1-A2

Method BLASTX
NCBI GI 94063747
BLAST score 491
E value 2.0e-66
Match length 248
% identity 58

NCBI Description (AC005851) hypothetical protein [Arabidopsis thaliana]

Seq. No. 3760

Contig ID 2799 1.R1040

5'-most EST LIB3170-046-Q1-J1-G2

Method BLASTX
NCBI GI g3341443
BLAST score 783
E value 2.0e-83
Match length 252
% identity 61

NCBI Description (AJ223074) acid phosphatase [Glycine max]

Seq. No. 3761

Contig ID 2801 1.R1040

5'-most EST LIB3049-029-Q1-E1-A12

Seq. No. 3762

Contig ID 2801_3.R1040 5'-most EST fde700871424.h1

Method BLASTX
NCBI GI g4454459
BLAST score 362
E value 2.0e-34
Match length 135
% identity 62

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No. 3763

Contig ID 2801_4.R1040 5'-most EST pmv700893892.h1

Method BLASTX
NCBI GI g4454459
BLAST score 228
E value 8.0e-19
Match length 77
% identity 58

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]



Seq. No. 3764

Contig ID 2807 1.R1040

5'-most EST LIB3109-047-Q1-K1-D2

Method BLASTN
NCBI GI g4519195
BLAST score 35
E value 6.0e-10
Match length 43
% identity 95

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, Pl clone:

MQC12, complete sequence

Seq. No. 3765

Contig ID 2809 1.R1040

5'-most EST LIB3056-006-Q1-N1-A4

Method BLASTX
NCBI GI g3367524
BLAST score 403
E value 6.0e-39
Match length 193
% identity 56

NCBI Description (AC004392) F8K4.12 [Arabidopsis thaliana]

Seq. No. 3766

Contig ID 2809 2.R1040

5'-most EST LIB3051-009-Q1-E1-A10

Seq. No. 3767

Contig ID 2810_1.R1040 5'-most EST wvk700681381.h2

Seq. No. 3768

Contig ID 2811 1.R1040

5'-most EST LIB $3\overline{0}28-039-Q1-B1-G12$

Method BLASTX
NCBI GI g1766046
BLAST score 738
E value 2.0e-78
Match length 153
% identity 90

NCBI Description (U81993) NAD+ dependent isocitrate dehydrogenase subunit 1

[Arabidopsis thaliana]

Seq. No. 3769

Contig ID 2812_1.R1040

5'-most EST LIB3039-034-Q1-E1-D5

Method BLASTX
NCBI GI g3004565
BLAST score 333
E value 1.0e-30
Match length 123
% identity 58

NCBI Description (AC003673) putative protein kinase [Arabidopsis thaliana]

Seq. No. 3770

Contig ID 2822 1.R1040

5'-most EST LIB3028-039-Q1-B1-D8



```
BLASTX
Method
                  g2651310
NCBI GI
                  193
BLAST score
                   6.0e-26
E value
                  152
Match length
                   39
% identity
                   (AC002336) putative PTR2-B peptide transporter [Arabidopsis
NCBI Description
                  thaliana]
                   3771
Seq. No.
                   2825 1.R1040
Contig ID
                   kl1701207614.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2827637
                   370
BLAST score
                   3.0e-35
E value
                   107
Match length
% identity
                   63
NCBI Description (AL021636) putative protein [Arabidopsis thaliana]
                   3772
Seq. No.
                   2828 1.R1040
Contig ID
                   qsv701054363.hl
5'-most EST
                   3773
Seq. No.
                   2828 2.R1040
Contig ID
                   LIB3072-017-Q1-E1-C2
5'-most EST
                   3774
Seq. No.
                   2829 1.R1040
Contig ID
                   jC-gmst02400018f02a1
5'-most EST
                   BLASTX
Method
                   g2244876
NCBI GI
                   1168
BLAST score
                   1.0e-128
E value
                   286
Match length
                   74
% identity
                   (Z97338) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3775
Seq. No.
                   2831 1.R1040
Contig ID
5'-most EST
                   leu701152319.hl
Method
                   BLASTX
NCBI GI
                   q2642443
BLAST score
                   826
                   2.0e-88
E value
                   371
Match length
% identity
                   (ACO02391) putative cytochrome P450 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3776
                   2832 1.R1040
Contig ID
5'-most EST
                   fC-gmle7000742258f1
Method
                   BLASTX
```

q1706822

1.0e-34

365

NCBI GI

E value

BLAST score



```
Match length
% identity
                  FLAVONOL SYNTHASE (FLS) >gi 421946 pir S33510 flavonol
NCBI Description
                  synthase - garden petunia >gi 311658 emb_CAA80264 (Z22543)
                  flavonol synthase [Petunia x hybrida]
                  3777
Seq. No.
                  2832 2.R1040
Contig ID
5'-most EST
                  LIB3028-039-Q1-B1-A6
                  BLASTX
Method
NCBI GI
                  g1706822
BLAST score
                  934
                   1.0e-101
E value
                  265
Match length
% identity
                   65
                  FLAVONOL SYNTHASE (FLS) >gi_421946_pir_S33510 flavonol
NCBI Description
                   synthase - garden petunia >gi 311658 emb CAA80264 (Z22543)
                   flavonol synthase [Petunia x hybrida]
Seq. No.
                   2837 1.R1040
Contig ID
                   asn701134785.h2
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3892051
                   1126
BLAST score
E value
                   1.0e-123
Match length
                   251
% identity
                   (AC002330) predicted NADH dehydrogenase 24 kD subunit
NCBI Description
                   [Arabidopsis thaliana]
                   3779
Seq. No.
                   2840 1.R1040
Contig ID
                   uC-gmrominsoy035c05b1
5'-most EST
Method
                   BLASTX
                   g2924792
NCBI GI
                   880
BLAST score
                   1.0e-94
E value
Match length
                   221
                   76
% identity
                   (ACO02334) similar to synaptobrevin [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3780
                   2841 1.R1040
Contig ID
                   zhf700958558,h1
5'-most EST
                   BLASTX
Method
```

q3150411 NCBI GI 1011 BLAST score 1.0e-110 E value 384 Match length % identity

(AC004165) hookless1-like protein [Arabidopsis thaliana] NCBI Description

Seq. No.

2842 1.R1040 Contig ID zzp700831646.hl 5'-most EST

BLASTN Method



```
g3176097
NCBI GI
BLAST score
                   346
                   0.0e + 00
E value
Match length
                   937
                   84
% identity
                   Medicago truncatula mRNA for annexin
NCBI Description
Seq. No.
                   2842 2.R1040
Contig ID
5'-most EST
                   rlr700900274.h1
Method
                   BLASTN
NCBI GI
                   q3176097
                   120
BLAST score
E value
                   9.0e-61
Match length
                   308
                   85
% identity
NCBI Description Medicago truncatula mRNA for annexin
                   3783
Seq. No.
                   2844 1.R1040
                   k117\overline{0}1206547.h1
                   BLASTN
                   g1370165
```

```
Contig ID
5'-most EST
Method
NCBI GI
BLAST score
                   558
E value
                   0.0e + 00
                   877
Match length
                   92
```

3784

% identity L.japonicus mRNA for small GTP-binding protein, RAB1C NCBI Description

Seq. No. 2844 3.R1040 Contig ID hrw701058664.hl 5'-most EST BLASTN Method g1370165 NCBI GI 157 BLAST score 6.0e-83 E value 296 Match length 95 % identity

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB1C

3785 Seq. No. 2846 1.R1040 Contig ID 5'-most EST trc700564010.hl Method BLASTX q3551954 NCBI GI 584 BLAST score

5.0e-60 E value 266 Match length % identity 43

(AF082030) senescence-associated protein 5 [Hemerocallis NCBI Description

hybrid cultivar]

3786 Seq. No.

2847 1.R1040 Contig ID 5'-most EST xpa700796473.h1

3787 Seq. No.



```
2848 1.R1040
Contig ID
                   kmv700743979.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2244866
                   912
BLAST score
E value
                   2.0e-98
                   249
Match length
                   68
% identity
                  (Z97337) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3788
Seq. No.
                   2849 1.R1040
Contig ID
                   jC-qmro02800031c11a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2660670
BLAST score
                   1056
                   1.0e-115
E value
                   282
Match length
                   73
% identity
                   (AC002342) putative Cu2+-transporting ATPase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   3789
                   2851 1.R1040
Contig ID
                   LIB3\overline{0}56-008-Q1-N1-A11
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2464865
BLAST score
                   1848
                   0.0e+00
E value
                   516
Match length
                   66
% identity
                   (Z99707) pectinesterase like protein [Arabidopsis thaliana]
NCBI Description
                   3790
Seq. No.
                   2851 2.R1040
Contig ID
                   LIB3107-059-Q1-K1-D7
5'-most EST
                   BLASTX
Method
                   g2464865
NCBI GI
BLAST score
                   372
                   2.0e-35
E value
                   104
Match length
                   64
% identity
                   (Z99707) pectinesterase like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    3791
                    2851 3.R1040
Contig ID
                   qsv701051641.h1
5'-most EST
                   BLASTX
Method
                   q2464865
NCBI GI
                    274
BLAST score
                    2.0e-24
E value
Match length
                    66
 % identity
                   (Z99707) pectinesterase like protein [Arabidopsis thaliana]
NCBI Description
```

Seq. No. 3792



```
5'-most EST
                  leu701147408.h1
                  BLASTX
Method
                  g2464865
NCBI GI
                  210
BLAST score
                   8.0e-17
E value
                  55
Match length
% identity
                  (Z99707) pectinesterase like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2852 1.R1040
Contig ID
5'-most EST
                  LIB3051-059-Q1-K2-F11
                  BLASTX
Method
NCBI GI
                   q3132479
BLAST score
                   2193
                   0.0e+00
E value
                   572
Match length
                   54
% identity
                   (AC003096) multidrug resistance-associated protein, AtMRP2
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   3794
                   2852 2.R1040
Contig ID
                   jC-qmf102220104d08d1
5'-most EST
                   3795
Seq. No.
                   2855 1.R1040
Contig ID
5'-most EST
                   dpv701098968.h1
                   BLASTX
Method
                   g3250695
NCBI GI
BLAST score
                   541
                   8.0e-55
E value
                   309
Match length
                   64
% identity
                   (AL024486) putative protein [Arabidopsis thaliana]
NCBI Description
                   3796
Seq. No.
Contig ID
                   2855 2.R1040
                   g4260400
5'-most EST
                   BLASTX
Method
                   g3250695
NCBI GI
BLAST score
                   163
E value
                   5.0e-11
Match length
                   70
% identity
                   (AL024486) putative protein [Arabidopsis thaliana]
NCBI Description
                   3797
Seq. No.
Contig ID
                   2857 1.R1040
                   LIB3050-020-Q1-K1-F6
5'-most EST
Method
                   BLASTN
                   g3204124
NCBI GI
BLAST score
                   105
E value
                   1.0e-51
                   260
Match length
                   85
% identity
```

NCBI Description Cicer arietinum mRNA for putative Pi starvation-induced

NCBI Description



protein

```
3798
Seq. No.
                  2857 2.R1040
Contig ID
                  LIB3028-038-Q1-B1-D9
5'-most EST
                   3799
Seq. No.
                   2858 1.R1040
Contig ID
5'-most EST
                  LIB3106-033-Q1-K1-E2
                   BLASTX
Method
NCBI GI
                   q2558512
BLAST score
                   159
                   2.0e-10
E value
Match length
                   91
% identity
                   41
                   (AJ002020) proton pump interactor [Arabidopsis thaliana]
NCBI Description
                   3800
Seq. No.
                   2861 1.R1040
Contig ID
                   LIB3170-057-Q1-K1-A2
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2827536
BLAST score
                   302
E value
                   2.0e-26
Match length
                   162
% identity
                   53
                   (AL021633) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3801
Seq. No.
                   2861 2.R1040
Contig ID
                   jC-gmle01810093a10d1
5'-most EST
                   3802
Seq. No.
                   2861 3.R1040
Contig ID
5'-most EST
                   jC-gmle01810030f03a2
                   3803
Seq. No.
                   2861 4.R1040
Contig ID
5'-most EST
                   LIB3028-038-Q1-B1-E5
                   BLASTX
Method
                   g2827536
NCBI GI
BLAST score
                   184
                   2.0e-13
E value
Match length
                   45
% identity
                   (AL021633) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   3804
Seq. No.
                   2864 1.R1040
Contig ID
                   rlr700896672.hl
5'-most EST
                   BLASTX
Method
                   q4263787
NCBI GI
BLAST score
                   883
                    4.0e-95
E value
Match length
                   301
 % identity
                   (AC006068) unknown protein [Arabidopsis thaliana]
```



```
Seq. No.
                   3805
                   2864 2.R1040
Contig ID
                   LIB3049-010-Q1-E1-G3
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4263787
                   292
BLAST score
E value
                   4.0e-26
Match length
                   75
% identity
                   76
                   (AC006068) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3806
Seq. No.
                   2864 3.R1040
Contig ID
                   bth700846001.h1
5'-most EST
                   3807
Seq. No.
                   2865 1.R1040
Contig ID
5'-most EST
                   leu701154015.h1
                   3808
Seq. No.
                   2869 1.R1040
Contig ID
                   LIB3049-045-Q1-E1-D6
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3158476
BLAST score
                   1195
                   1.0e-131
E value
                   285
Match length
% identity
                   80
                   (AF067185) aquaporin 2 [Samanea saman]
NCBI Description
                   3809
Seq. No.
                   2871 1.R1040
Contig ID
                   jC-qmf102220069d08a1
5'-most EST
                   3810
Seq. No.
                   2873 1.R1040
Contig ID
                   leu7\overline{0}1149165.h1
5'-most EST
                   BLASTX
Method
                   g1435021
NCBI GI
                   662
BLAST score
                   5.0e-69
E value
                   284
Match length
                   50
% identity
                   (D26575) DNA-binding protein [Daucus carota]
NCBI Description
                    3811
Seq. No.
                    2873 2.R1040
Contig ID
5'-most EST
                   uC-gmropic040b12b1
Seq. No.
                    3812
                    2873 4.R1040
Contig ID
                   LIB3139-005-P1-N1-A11
5'-most EST
```

Method BLASTN
NCBI GI g1161574
RLAST score 35

BLAST score 35 E value 3.0e-10



```
51
Match length
% identity
                  92
NCBI Description L.esculentum mRNA for homeobox protein
                  3813
Seq. No.
                  2878 1.R1040
Contig ID
                  zhf700965057.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3641868
BLAST score
                  320
                  3.0e-29
E value
                  108
Match length
                  60
% identity
                  (AJ011012) hypothetical protein [Cicer arietinum]
NCBI Description
                  3814
Seq. No.
                  2879 1.R1040
Contig ID
5'-most EST
                  LIB3106-104-Q1-K1-C9
Method
                  BLASTX
NCBI GI
                   g1854386
BLAST score
                   980
E value
                  1.0e-106
Match length
                  269
% identity
                   68
                   (AB001375) similar to soluble NSF attachment protein [Vitis
NCBI Description
                  vinifera]
                   3815
Seq. No.
                   2879 2.R1040
Contig ID
                  LIB3139-013-P1-N1-B8
5'-most EST
Method
                  BLASTN
NCBI GI
                   g1854385
BLAST score
                   56
                   1.0e-22
E value
Match length
                  144
% identity
                   85
                  Vitis vinifera mRNA for soluble NSF attachment protein
NCBI Description
                  homologue, complete cds
                   3816
Seq. No.
Contig ID
                   2885 1.R1040
5'-most EST
                   LIB3139-093-P1-N1-F2
Method
                   BLASTX
NCBI GI
                   q2623295
BLAST score
                   457
E value
                   2.0e-75
Match length
                   219
% identity
NCBI Description
                   (AC002409) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   3817
```

Contig ID 2886 1.R1040

5'-most EST LIB3109-037-Q1-K1-E6

Method BLASTX NCBI GI q280401 BLAST score 1497 1.0e-167 E value



Match length 389 % identity 76

NCBI Description H+-transporting ATPase (EC 3.6.1.35) - curled-leaved

tobacco >gi 170206 (M27888) H+-translocating ATPase

[Nicotiana plumbaginifolia]

Seq. No. 3818

Contig ID 2887_1.R1040 5'-most EST epx701106441.h1

Method BLASTX
NCBI GI g4150963
BLAST score 565
E value 5.0e-58
Match length 169
% identity 60

NCBI Description (Y18620) DsPTP1 protein [Arabidopsis thaliana]

Seq. No. 3819

Contig ID 2887_2.R1040 5'-most EST asn701141361.h1

Seq. No. 3820

Contig ID 2888_1.R1040 5'-most EST jsh701070307.h1

Method BLASTX
NCBI GI g4510373
BLAST score 628
E value 3.0e-65
Match length 191
% identity 57

NCBI Description (AC007017) putative harpin-induced protein [Arabidopsis

thaliana]

Seq. No. 3821

Contig ID 2888_2.R1040 5'-most EST jex700906012.h1

Method BLASTX
NCBI GI 94510373
BLAST score 293
E value 2.0e-26
Match length 88
% identity 60

NCBI Description (AC007017) putative harpin-induced protein [Arabidopsis

thaliana]

Seq. No. 3822

Contig ID 2892_1.R1040

5'-most EST LIB3028-038-Q1-B1-A6

Seq. No. 3823

Contig ID 2895_1.R1040 5'-most EST pmv700894421.h1

Method BLASTX
NCBI GI g3702332
BLAST score 306
E value 5.0e-63
Match length 248

Match length

% identity

86 50



```
% identity
                  (AC005397) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  3824
                  2895 2.R1040
Contig ID
5'-most EST
                  LIB3040-061-Q1-E11-E4
Method
                  BLASTX
NCBI GI
                  g4454032
BLAST score
                  257
                  3.0e-22
E value
Match length
                  89
                  56
% identity
NCBI Description (AL035394) putative protein [Arabidopsis thaliana]
Seq. No.
                  3825
                  2897 1.R1040
Contig ID
                  LIB3051-040-Q1-K1-F12
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2921823
BLAST score
                   327
                   6.0e-30
E value
Match length
                  144
                   55
% identity
                  (AF046934) shoot-forming PKSF1 [Paulownia kawakamii]
NCBI Description
                   3826
Seq. No.
                   2897 2.R1040
Contig ID
                   LIB3028-038-Q1-B1-B5
5'-most EST
                   BLASTX
Method
                   g2246376
NCBI GI
                   171
BLAST score
                   4.0e-12
E value
                   62
Match length
% identity
                  (Z86093) b-Zip DNA binding protein [Arabidopsis thaliana]
NCBI Description
                   3827
Seq. No.
                   2898 1.R1040
Contig ID
                   jex700905254.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2961384
BLAST score
                   383
                   1.0e-36
E value
Match length
                   140
% identity
                   (AL022141) aldehyde dehydrogenase like protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   3828
Contig ID
                   2898 2.R1040
                   LIB3028-038-Q1-B1-B6
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2961384
BLAST score
                   238
                   7.0e-20
E value
```



NCBI Description (AL022141) aldehyde dehydrogenase like protein [Arabidopsis thaliana]

Seq. No. 3829

Contig ID 2899_1.R1040

5'-most EST LIB3170-002-Q1-K1-A5

Method BLASTN
NCBI GI g498907
BLAST score 170
E value 2.0e-90
Match length 382
% identity 88

NCBI Description Pisum sativum ribosomal protein L34 homolog (RPL34) mRNA,

complete cds

Seq. No. 3830

Contig ID 2900_1.R1040

5'-most EST uC-gmflminsoy002d01b1

Method BLASTX
NCBI GI g1177320
BLAST score 487
E value 1.0e-48
Match length 176
% identity 53

NCBI Description (X89891) EFA27 for EF hand, abscisic acid, 27kD [Oryza

sativa]

Seq. No. 3831

Contig ID 2900_3.R1040

5'-most EST LIB3051-048-Q1-K1-D5

Method BLASTX
NCBI GI g2270994
BLAST score 121
E value 1.0e-13
Match length 57
% identity 63

NCBI Description (AF004809) Ca+2-binding EF hand protein [Glycine max]

Seq. No. 3832

Contig ID 2901_1.R1040

5'-most EST LIB3093-022-Q1-K1-G7

Seq. No. 3833

Contig ID 2901_2.R1040

5'-most EST LIB3049-006-Q1-E1-E10

Method BLASTX
NCBI GI g114420
BLAST score 2306
E value 0.0e+00
Match length 513
% identity 89

NCBI Description ATP SYNTHASE BETA CHAIN, MITOCHONDRIAL PRECURSOR

>gi_100882_pir__S11491 H+-transporting ATP synthase (EC

3.6.1.34) beta chain, mitochondrial - maize

>gi_22173_emb_CAA38140_ (X54233) ATPase F1 subunit protein
[Zea_mays] >gi_897618 (M36087) F-1-ATPase subunit 2 [Zea

mays]



```
3834
Seq. No.
                  2901 4.R1040
Contig ID
                  asn701142701.hl
5'-most EST
                  BLASTN
Method
                  g2204063
NCBI GI
                  67
BLAST score
E value
                  2.0e-29
                  131
Match length
                   92
% identity
NCBI Description Pisum sativum mRNA for F1 ATPase, complete cds
                   3835
Seq. No.
                   2901 5.R1040
Contig ID
                   zsg701119274.hl
5'-most EST
                   BLASTN
Method
                   g2204063
NCBI GI
BLAST score
                   34
                   1.0e-09
E value
Match length
                   70
                   87
% identity
                   Pisum sativum mRNA for F1 ATPase, complete cds
NCBI Description
                   3836
Seq. No.
                   2906 1.R1040
Contig ID
                   fua701042256.hl
5'-most EST
                   BLASTX
Method
                   g1688074
NCBI GI
                   251
BLAST score
                   4.0e-21
E value
                   193
Match length
                   35
% identity
                   (U46570) tetratricopeptide repeat protein [Homo sapiens]
NCBI Description
                   >gi_4507711_ref_NP_003305.1_pTTC1_ tetratricopeptide repeat
                   domain
                   3837
Seq. No.
                   2906 2.R1040
Contig ID
                   hrw701059481.hl
5'-most EST
                   3838
Seq. No.
                   2909 1.R1040
Contig ID
5'-most EST
                   epx701108917.h1
Method
                   BLASTX
                   q1172494
NCBI GI
BLAST score
                   161
                   2.0e-10
E value
                   113
Match length
% identity
                   30
                   PTERIN-4-ALPHA-CARBINOLAMINE DEHYDRATASE (PHS)
NCBI Description
                    (4-ALPHA-HYDROXY-TETRAHYDROPTERIN DEHYDRATASE)
                    (PHENYLALANINE HYDROXYLASE-STIMULATING PROTEIN) (PCD)
                    >gi_476742 (M88627) phenylalanine hydroxylase [Pseudomonas
```

3839 2909 3.R1040 Contig ID

Seq. No.

aeruginosa]



```
LIB3028-037-Q1-B1-H11
5'-most EST
                  3840
Seq. No.
                  2910 1.R1040
Contig ID
                  k117\overline{0}1211587.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3183568
                  634
BLAST score
                  1.0e-65
E value
                  340
Match length
% identity
                  41
                  HYPOTHETICAL 45.1 KD PROTEIN IN CDD-MGLC INTERGENIC REGION
NCBI Description
Seq. No.
                  2910 2.R1040
Contig ID
                  trc700563552.hl
5'-most EST
                   3842
Seq. No.
                   2912 1.R1040
Contig ID
                  LIB3049-051-Q1-E1-D8
5'-most EST
                  BLASTX
Method
                   g3367593
NCBI GI
BLAST score
                   512
                   8.0e-52
E value
                   125
Match length
                   78
% identity
                   (AL031135) putative protein [Arabidopsis thaliana]
NCBI Description
                   >gi 3805841 emb_CAA21461_ (AL031986) putative protein
                   [Arabidopsis thaliana]
                   3843
Seq. No.
                   2912 2.R1040
Contig ID
5'-most EST
                   LIB3109-035-Q1-K1-A5
Method
                   BLASTX
                   q3367593
NCBI GI
BLAST score
                   195
                   2.0e-22
E value
Match length
                   81
                   73
% identity
                   (AL031135) putative protein [Arabidopsis thaliana]
NCBI Description
                   >gi 3805841 emb CAA21461 (AL031986) putative protein
                   [Arabidopsis thaliana]
Seq. No.
                   3844
                   2913 1.R1040
Contig ID
                   LIB3028-037-Q1-B1-H6
5'-most EST
Method
                   BLASTX
NCBI GI
                   q98837
BLAST score
                   177
                   7.0e-13
E value
                   106
Match length
% identity
                   36
                   enantiomer-selective amidase - Rhodococcus sp >gi_152052
NCBI Description
                   (M74531) enantiomerase-selective amidase [Rhodococcus sp.]
```

3845

2914 1.R1040

Seq. No.

Contig ID

5'-most EST

Method NCBI GI



```
epx701105154.hl
5'-most EST
                  BLASTX
Method
                  g1209756
NCBI GI
                  173
BLAST score
                   1.0e-32
E value
                   119
Match length
                   68
% identity
                   (U43629) integral membrane protein [Beta vulgaris]
NCBI Description
Seq. No.
                   3846
                   2914 2.R1040
Contig ID
                   pxt700946327.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1209756
                   177
BLAST score
                   1.0e-12
E value
                   134
Match length
                   49
% identity
NCBI Description (U43629) integral membrane protein [Beta vulgaris]
                   3847
Seq. No.
                   2915 1.R1040
Contig ID
                   LIB3055-013-Q1-N1-G8
5'-most EST
                   BLASTX
Method
                   g2924507
NCBI GI
                   809
BLAST score
                   2.0e-86
E value
                   178
Match length
% identity
                   83
                   (AL022023) cyclophilin - like protein [Arabidopsis
NCBI Description
                   thaliana]
                   3848
Seq. No.
                   2918 1.R1040
Contig ID
5'-most EST
                   LIB3039-013-Q1-E1-H3
                   3849
Seq. No.
Contig ID
                   2918 2.R1040
                   LIB3049-051-Q1-E1-H3
5'-most EST
                   3850
Seq. No.
Contig ID
                   2919 1.R1040
5'-most EST
                   qsv701048547.hl
Method
                   BLASTX
                   q3122703
NCBI GI
BLAST score
                   586
                   2.0e-60
E value
Match length
                   138
% identity
                   84
                   60S RIBOSOMAL PROTEIN L23A >gi_2641201 (AF031542) ribosomal
NCBI Description
                   protein L23a [Fritillaria agrestis]
Seq. No.
                   3851
Contig ID
                   2920 1.R1040
```

jC-gmle01810053a11a1

BLASTX

g1173198



```
BLAST score
                  2.0e-77
E value
Match length
                  151
% identity
                   95
                  40S RIBOSOMAL PROTEIN S13 >gi_480095_pir__S36423 ribosomal
NCBI Description
                  protein S13.e - garden pea >gi_396639_emb_CAA80974_
                   (Z25509) ribosomal protein S13 [Pisum sativum]
Seq. No.
                   3852
                   2920 2.R1040
Contig ID
                  LIB3138-033-Q1-N1-B6
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3450842
BLAST score
                   415
E value
                   2.0e-40
Match length
                   212
                   45
% identity
                   (AF080436) mitogen activated protein kinase kinase [Oryza
NCBI Description
                   sativa]
                   3853
Seq. No.
                   2920 3.R1040
Contig ID
                   LIB3094-002-Q1-K1-A7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1173198
                   723
BLAST score
                   2.0e-76
E value
Match length
                   151
% identity
                   93
                   40S RIBOSOMAL PROTEIN S13 >gi_480095_pir__S36423 ribosomal
NCBI Description
                   protein S13.e - garden pea >gi_396639 emb CAA80974
                   (Z25509) ribosomal protein S13 [Pisum satīvum]
                   3854
Seq. No.
                   2920 5.R1040
Contig ID
                   2DC-\overline{0}1-Q1-E1-E7
5'-most EST
                   BLASTN
Method
                   g3193311
NCBI GI
BLAST score
                   43
                   7.0e-15
E value
                   71
Match length
                   90
% identity
                   Arabidopsis thaliana BAC F6N15
NCBI Description
Seq. No.
                   3855
                   2920 8.R1040
Contig ID
                   taw700657103.h1
5'-most EST
                   3856
Seq. No.
                   2923 1.R1040
Contig ID
                   ncj700976784.hl
```

Contig ID 2923_1.R1040
5'-most EST ncj700976784.
Method BLASTX
NCBI GI g3046700
BLAST score 337

E value 1.0e-31 Match length 91 70



NCBI Description (AJ005261) cytidine deaminase [Arabidopsis thaliana]

>gi_3093276_emb_CAA06671 (AJ005687) cytidine deaminase
[Arabidopsis thaliana] >gi_4191787 (AC005917) putative

cytidine deaminase [Arabidopsis thaliana]

Seq. No. 3857

Contig ID 2926_1.R1040 5'-most EST kl1701204486.h2

Seq. No. 3858

Contig ID 2926_2.R1040

5'-most EST LIB3028-037-Q1-B1-G1

Seq. No. 3859

Contig ID 2929 1.R1040

5'-most EST jC-gmf102220063f08a1

Method BLASTX
NCBI GI g3775993
BLAST score 1425
E value 0.0e+00
Match length 424

Match length 424 % identity 80

NCBI Description (AJ010460) RNA helicase [Arabidopsis thaliana]

Seq. No. 3860

Contig ID 2929 2.R1040

5'-most EST uC-gmrominsoy122f10b1

Seq. No. 3861

Contig ID 2929 3.R1040

5'-most EST LIB3 $\overline{0}$ 51-090-Q1-K1-F5

Seq. No. 3862

Contig ID 2929 4.R1040

5'-most EST LIB3051-084-Q1-K1-C9

Seq. No. 3863

Contig ID 2930 1.R1040

5'-most EST LIB3040-061-Q1-E1-A11

Method BLASTN
NCBI GI g20761
BLAST score 58
E value 2.0e-23
Match length 225
% identity 84

NCBI Description Pea mRNA for H1 histone

Seq. No. 3864

Contig ID 2930_2.R1040

5'-most EST LIB3049-056-Q1-E1-G12

Method BLASTN
NCBI GI g20761
BLAST score 64
E value 6.0e-27
Match length 232

Match length 232 % identity 87

A MARINE TO THE REAL PROPERTY.

NCBI Description Pea mRNA for H1 histone



```
3865
Seq. No.
                  2930 4.R1040
Contig ID
5'-most EST
                  LIB3170-030-Q1-K1-H6
                  3866
Seq. No.
                  2931 1.R1040
Contig ID
                  jC-gmst02400055b09a1
5'-most EST
                  BLASTX
Method
                  g1800147
NCBI GI
                  280
BLAST score
                  8.0e-25
E value
                  91
Match length
                   68
% identity
NCBI Description (U83655) membrane associated protein [Arabidopsis thaliana]
                   3867
Seq. No.
                   2935 1.R1040
Contig ID
                   jC-gmst02400026e07a1
5'-most EST
                   3868
Seq. No.
                   2939 1.R1040
Contig ID
5'-most EST
                   leu701149477.hl
                   BLASTX
Method
                   q4432835
NCBI GI
                   210
BLAST score
                   1.0e-16
E value
                   80
Match length
                   47
% identity
                   (AC006283) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3869
Seq. No.
                   2939 2.R1040
Contig ID
                   LIB3106-094-Q1-K1-D11
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4432835
                   249
BLAST score
                   3.0e-21
E value
                   95
Match length
                   48
% identity
                   (AC006283) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3870
Seq. No.
                   2943 1.R1040
Contig ID
                   LIB3028-037-Q1-B1-E4
5'-most EST
Method
                   BLASTX
                   g461498
NCBI GI
                   874
BLAST score
                   4.0e-94
E value
                   221
Match length
                   73
 % identity
                   ALANINE AMINOTRANSFERASE 2 (GPT) (GLUTAMIC--PYRUVIC
NCBI Description
                   TRANSAMINASE 2) (GLUTAMIC--ALANINE TRANSAMINASE 2)
                    (ALAAT-2) >gi_320619_pir__S28429 alanine transaminase (EC
                   2.6.1.2) - proso millet \overline{>}gi_296204_emb_CAA49199_ (X69421)
```

alanine aminotransferase [Panicum miliaceum]



```
3871
Seq. No.
                  2945 1.R1040
Contig ID
                  uC-gmflminsoy099g11b1
5'-most EST
                  3872
Seq. No.
                  2946 1.R1040
Contig ID
                  LIB3109-005-Q1-K1-C3
5'-most EST
Method
                  BLASTX
NCBI GI
                   q2832686
BLAST score
                   214
                   4.0e-17
E value
                   69
Match length
% identity
                   (AL021712) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   2947 1.R1040
Contig ID
                   uC-gmrominsoy263h01b1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4063746
                   700
BLAST score
                   2.0e-73
E value
                   506
Match length
% identity
                   33
NCBI Description (AC005851) nodulin-like protein [Arabidopsis thaliana]
                   3874
Seq. No.
                   2947 2.R1040
Contig ID
                   uC-gmrominsoy048a04b1
5'-most EST
                   3875
Seq. No.
                   2947 3.R1040
Contig ID
                   uC-gmrominsoy165e07b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4455155
                   350
BLAST score
                   4.0e-33
E value
                   115
Match length
% identity
                   65
                   (AL022023) EF-1 alpha - like protein (fragment)
NCBI Description
                   [Arabidopsis thaliana]
                   3876
Seq. No.
                   2948 1.R1040
Contig ID
                   pmv700893325.h1
5'-most EST
                   3877
Seq. No.
                   2948 2.R1040
Contig ID
5'-most EST
                   jC-gmle01810052g12a1
```

3878 Seq. No.

2950 1.R1040 Contig ID zhf700963838.h1 5'-most EST

Seq. No. 3879

2950 2.R1040 Contig ID

jC-qmle01810033d04a2 5'-most EST



```
3880
Seq. No.
                  2950 3.R1040
Contig ID
                  jC-gmst02400050f12d1
5'-most EST
                  3881
Seq. No.
                  2950 5.R1040
Contig ID
                   jex700904169.h1
5'-most EST
                  BLASTN
Method
                  g312300
NCBI GI
BLAST score
                   36
                   2.0e-10
E value
                   52
Match length
                   92
% identity
NCBI Description G.max cDNA for glutamine synthetase (3' region)
                   3882
Seq. No.
                   2950 6.R1040
Contig ID
                   LIB3051-109-Q1-K1-A2
5'-most EST
                   3883
Seq. No.
                   2951 1.R1040
Contig ID
                   LIB3056-011-Q1-N1-C6
5'-most EST
Method
                   BLASTX
                   g3252856
NCBI GI
                   2200
BLAST score
                   0.0e + 00
E value
                   501
Match length
% identity
                   (AF020425) glutamate decarboxylase isozyme 1 [Nicotiana
NCBI Description
                   tabacum]
                   3884
Seq. No.
                   2951 2.R1040
Contig ID
                   fC-qmse700654980f1
5'-most EST
Method
                   BLASTX
                   g4510342
NCBI GI
BLAST score
                   516
                   2.0e-52
E value
                   133
Match length
                   73
% identity
                   (AC006921) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
                   3885
Seq. No.
                   2951 3.R1040
Contig ID
                   LIB3051-091-Q1-K1-E6
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4510342
```

Method BLASTX
NCBI GI 94510342
BLAST score 355
E value 1.0e-33
Match length 104
% identity 66

NCBI Description (AC006921) putative serine/threonine protein kinase

[Arabidopsis thaliana]

Seq. No. 3886



```
2952 1.R1040
Contig ID
5'-most EST
                  LIB3170-006-Q1-J1-A6
Seq. No.
                   3887
Contig ID
                  2953 1.R1040
                  uC-gmrominsoy169g02b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1304599
BLAST score
                   231
                   5.0e-19
E value
Match length
                   178
                   32
% identity
NCBI Description (U41315) ZNF127-Xp [Homo sapiens]
Seq. No.
                   3888
                   2954 1.R1040
Contig ID
5'-most EST
                  uC-gmropic016e12b1
Method
                  BLASTX
NCBI GI
                  g3367523
BLAST score
                   419
                   5.0e-41
E value
Match length
                  121
                   63
% identity
                  (AC004392) ESTs gb AA728658 and gb N95943 come from this
NCBI Description
                  gene. [Arabidopsis thaliana]
Seq. No.
                   3889
Contig ID
                  2956 1.R1040
5'-most EST
                  LIB3028-037-Q1-B1-C3
Method
                  BLASTX
NCBI GI
                  g1504008
BLAST score
                   479
E value
                   2.0e-95
Match length
                   427
                   48
% identity
                  (D86967) Containing ATP/GTP-binding site motif A(P-loop):
NCBI Description
                   Similar to C.elegans protein(P1:CEC47E128); Similar to Mouse
                   alpha-mannosidase(P1:B54407) [Homo sapiens]
Seq. No.
                   3890
Contig ID
                   2960 1.R1040
5'-most EST
                   seb700651089.h1
Method
                   BLASTX
NCBI GI
                   g2244806
BLAST score
                   415
E value
                   1.0e-40
                  158
Match length
                   52
% identity
NCBI Description (Z97336) hypothetical protein [Arabidopsis thaliana]
```

Seq. No.

3891

Contig ID 2961 1.R1040

5'-most EST LIB3170-015-Q1-K1-F2

Method BLASTX
NCBI GI g3334333
BLAST score 447
E value 6.0e-61



Match length 155 % identity 74

NCBI Description SUPEROXIDE DISMUTASE-2 (CU-ZN) >gi_2660798 (AF034832)

cytosolic copper/zinc superoxide dismutase

[Mesembryanthemum crystallinum]

Seq. No. 3892

Contig ID 2963_1.R1040

5'-most EST jC-gmf102220093c12a1

Method BLASTX
NCBI GI g2979547
BLAST score 343
E value 5.0e-36
Match length 202
% identity 41

NCBI Description (AC003680) putative 7-ethoxycoumarin O-deethylase

[Arabidopsis thaliana]

Seq. No. 3893

Contig ID 2964_1.R1040

5'-most EST LIB3030-004-Q1-B1-D6

Seq. No. 3894

Contig ID 2965_1.R1040 5'-most EST hyd700726944.h1

Method BLASTX
NCBI GI g2827715
BLAST score 1946
E value 0.0e+00
Match length 482
% identity 71

NCBI Description (AL021684) receptor protein kinase - like protein

[Arabidopsis thaliana]

Seq. No. 3895

Contig ID 2965 2.R1040

5'-most EST fC-gmle700870555g2

Method BLASTX
NCBI GI g2827715
BLAST score 551
E value 2.0e-56
Match length 120
% identity 88

NCBI Description (AL021684) receptor protein kinase - like protein

[Arabidopsis thaliana]

Seq. No. 3896

Contig ID 2966_1.R1040
5'-most EST g5677385
Method BLASTX
NCBI GI g2654226
BLAST score 626
E value 3.0e-65
Match length 155
% identity 79

NCBI Description (AJ003069) aminoacyl-t-RNA synthetase [Arabidopsis

thaliana]



```
3897
Seq. No.
                   2968 1.R1040
Contig ID
5'-most EST
                   zhf700963533.h1
                   BLASTX
Method
                   q4415912
NCBI GI
                   199
BLAST score
                   4.0e-15
E value
                   86
```

% identity NCBI Description (AC006282) putative protease [Arabidopsis thaliana]

3898 Seq. No.

Match length

2969 1.R1040 Contig ID

5'-most EST LIB3092-057-Q1-K1-B8

51

3899 Seq. No.

2969 2.R1040 Contig ID

LIB3092-057-Q1-K1-B10 5'-most EST

3900 Seq. No.

2969 3.R1040 Contig ID

5'-most EST LIB3051-035-Q1-K1-E9

3901 Seq. No.

2971 1.R1040 Contig ID

LIB3028-037-Q1-B1-A5 5'-most EST

3902 Seq. No.

2973 1.R1040 Contig ID

5'-most EST jC-gmst02400056g01a1

BLASTX Method g3413511 NCBI GI BLAST score 1302 E value 1.0e-144 Match length 272

% identity 89

(AJ000265) glucose-6-phosphate isomerase [Spinacia NCBI Description

oleracea]

3903 Seq. No.

2973 2.R1040 Contig ID $hyd7\overline{0}0730316.h1$ 5'-most EST

BLASTX Method g4220528 NCBI GI BLAST score 1486 1.0e-165 E value 380 Match length

% identity (AL035356) glucose-6-phosphate isomerase [Arabidopsis NCBI Description

thaliana]

3904 Seq. No.

2974 1.R1040 Contig ID 5'-most EST epx701105525.h1

3905 Seq. No.



```
2975 1.R1040
Contig ID
                  LIB3039-027-Q1-E1-B4
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2583134
                  258
BLAST score
                  3.0e-22
E value
Match length
                  124
                  50
% identity
                   (AC002387) putative proline-rich protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                   3906
                   2979 1.R1040
Contig ID
                   pxt700944651.h1
5'-most EST
                   BLASTX
Method
                   g3582340
NCBI GI
                   350
BLAST score
                   4.0e-33
E value
Match length
                   92
                   76
% identity
                   (AC005496) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3907
Seq. No.
                   2982 1.R1040
Contig ID
                   g4307500
5'-most EST
                   BLASTN
Method
                   g2351065
NCBI GI
BLAST score
                   34
                   2.0e-09
E value
Match length
                   164
                   84
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MHF15, complete sequence [Arabidopsis thaliana]
                   3908
Seq. No.
                   2983 1.R1040
Contig ID
                   zhf700962539.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2582799
                   115
BLAST score
                   1.0e-57
E value
                   353
Match length
                   88
% identity
                   M.sativa mRNA for protein phosphatase 2C
NCBI Description
                   3909
Seq. No.
                   2983 2.R1040
Contig ID
                   dpv701103002.hl
5'-most EST
Method
                   BLASTX
                   g2582800
NCBI GI
BLAST score
                   809
                   2.0e-86
E value
Match length
                   209
% identity
                   76
```

Seq. No. 3910

NCBI Description

(Y11607) protein phosphatase 2C [Medicago sativa]



Contig ID 2983_3.R1040 5'-most EST uC-gmropic020b03b1

Method BLASTN
NCBI GI g2582799
BLAST score 194

E value 1.0e-105 Match length 417 % identity 88

NCBI Description M.sativa mRNA for protein phosphatase 2C

Seq. No. 3911

Contig ID 2983_4.R1040

5'-most EST LIB3139-095-P1-N1-G11

Method BLASTN
NCBI GI g2582799
BLAST score 113
E value 1.0e-56
Match length 266
% identity 88

NCBI Description M.sativa mRNA for protein phosphatase 2C

Seq. No. 3912

Contig ID 2983_5.R1040 5'-most EST wrg700786175.h2

Method BLASTN
NCBI GI g2582799
BLAST score 43
E value 6.0e-15
Match length 67
% identity 91

NCBI Description M.sativa mRNA for protein phosphatase 2C

Seq. No. 3913

Contig ID 2983_6.R1040

5'-most EST uC-gmrominsoy208a02b1

Seq. No. 3914

Contig ID 2984 1.R1040

5'-most EST LIB3065-004-Q1-N1-H9

Seq. No. 3915

Contig ID 2988_1.R1040

5'-most EST LIB3028-036-Q1-B1-G8

Seq. No. 3916

Contig ID 2989 1.R1040

5'-most EST LIB3170-042-Q1-J1-B12

Method BLASTX
NCBI GI g1279640
BLAST score 666
E value 6.0e-70
Match length 153
% identity 80

NCBI Description (X92204) NAM [Petunia x hybrida]

Seq. No. 3917

Contig ID 2989 2.R1040



```
5'-most EST
                   kl1701211357.h1
Method
                   BLASTX
NCBI GI
                   q1944132
BLAST score
                   558
E value
                   3.0e-57
Match length
                   162
% identity
                   68
NCBI Description
                   (AB002560) CUC2 [Arabidopsis thaliana]
Seq. No.
                   3918
                   2992 1.R1040
Contig ID
5'-most EST
                   LIB3167-029-P1-K1-G2
Method
                   BLASTX
NCBI GI
                   q3282505
BLAST score
                   196
                   8.0e-15
E value
Match length
                   139
% identity
                   35
NCBI Description
                  (AF020786) polyphenol oxidase precursor [Prunus armeniaca]
Seq. No.
                   3919
Contig ID
                   2993 1.R1040
5'-most EST
                   jC-gmf102220139a10a1
Method
                   BLASTX
NCBI GI
                   q4314378
BLAST score
                   451
E value
                   1.0e-44
Match length
                   231
% identity
                   (AC006232) putative lipase [Arabidopsis thaliana]
NCBI Description
                   3920
Seq. No.
                   2998 1.R1040
Contig ID
5'-most EST
                   vzy700752164.hl
Method
                   BLASTX
NCBI GI
                   g1531758
BLAST score
                   751
                   6.0e-80
E value
Match length
                   182
% identity
                   79
NCBI Description
                   (X98772) AUX1 [Arabidopsis thaliana] >gi 3335360 (AC003028)
                   unknown protein [Arabidopsis thaliana]
Seq. No.
                   3921
                   2999 1.R1040
Contig ID
5'-most EST
                   pxt700944601.h1
Method
                   BLASTX
NCBI GI
                   g1518059
BLAST score
                   287
                   2.0e-25
E value
```

Match length 121 % identity 45

(U65650) blue-copper binging protein III [Arabidopsis NCBI Description

thaliana] >gi_3395770 (AF039404) uclacyanin 3 [Arabidopsis

thaliana]

Seq. No. 3922

Match length

% identity

139

68



```
3000 1.R1040
Contig ID
                  ncj700979531.h2
5'-most EST
Method
                  BLASTX
                  g3290209
NCBI GI
BLAST score
                  282
                  5.0e-25
E value
Match length
                  98
% identity
                  58
                  (U78947) MADS-box protein 1 [Malus domestica]
NCBI Description
                  3923
Seq. No.
                  3001 1.R1040
Contig ID
5'-most EST
                  jex700906677.h1
Method
                  BLASTX
                  g3928543
NCBI GI
BLAST score
                  1541
                  1.0e-172
E value
Match length
                  473
% identity
                  61
                   (AB016819) UDP-glucose glucosyltransferase [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  3924
                  3002 1.R1040
Contig ID
                   jex700909978.hl
5'-most EST
Method
                  BLASTX
                  g4263772
NCBI GI
BLAST score
                  517
                  2.0e-52
E value
                  162
Match length
% identity
                  (AC006218) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3925
                   3005 1.R1040
Contig ID
                   zzp700834539.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3746062
BLAST score
                   229
E value
                   1.0e-26
Match length
                  112
% identity
                  (AC005311) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3926
Contig ID
                   3008 1.R1040
5'-most EST
                   leu701156067.h1
                   3927
Seq. No.
Contig ID
                   3013 1.R1040
5'-most EST
                   LIB3092-055-Q1-K1-G7
Method
                   BLASTX
NCBI GI
                   g3415115
BLAST score
                   489
                   3.0e-49
E value
```



```
NCBI Description (AF081202) villin 2 [Arabidopsis thaliana]
                  3928
Seq. No.
                  3021 1.R1040
Contig ID
5'-most EST
                  LIB3106-113-Q1-K1-E8
                  BLASTX
Method
NCBI GI
                  q4510381
BLAST score
                  343
E value
                  1.0e-31
Match length
                  123
% identity
                  61
                  (AC007017) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                  3022 1.R1040
5'-most EST
                  LIB3107-013-Q1-K1-G9
Method
                  BLASTX
NCBI GI
                  g3913518
BLAST score
                  563
E value
                  2.0e-65
Match length
                  196
% identity
                  70
                  3'(2'),5'-BISPHOSPHATE NUCLEOTIDASE
NCBI Description
                   (3'(2'),5-BISPHOSPHONUCLEOSIDE 3'(2')-PHOSPHOHYDROLASE)
                   (DPNPASE) >gi_1103921 (U40433) 3'(2'),5'-bisphosphate
                  nucleotidase [Arabidopsis thaliana]
                  3930
Seq. No.
                  3022 2.R1040
Contig ID
                  wrg700790744.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3913518
                  291
BLAST score
                   4.0e-26
E value
                   91
Match length
                   66
% identity
                  3'(2'),5'-BISPHOSPHATE NUCLEOTIDASE
NCBI Description
                   (3'(2'),5-BISPHOSPHONUCLEOSIDE 3'(2')-PHOSPHOHYDROLASE)
                   (DPNPASE) >gi 1103921 (U40433) 3'(2'),5'-bisphosphate
                  nucleotidase [Arabidopsis thaliana]
                   3931
Seq. No.
Contig ID
                   3027 1.R1040
5'-most EST
                  bth700848545.h1
                   BLASTX
Method
                   q3128218
NCBI GI
BLAST score
                   617
E value
                   5.0e-64
Match length
                   204
% identity
                   (AC004077) putative end13 protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3932
```

Contig ID 3030 1.R1040

LIB3167-013-P1-K1-D5 5'-most EST

Method BLASTX NCBI GI q4325341

782



BLAST score 699 E value 8.0e-74 Match length 162 % identity

(AF128393) similar to the Drosophila DES-1 protein NCBI Description

(GB:X94180) [Arabidopsis thaliana]

3933 Seq. No.

Contig ID 3033 1.R1040 5'-most EST zhf700957311.h1

Method BLASTX NCBI GI q2979552 BLAST score 1169 1.0e-128 E value Match length 302 74 % identity

(AC003680) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No. 3934

3034 1.R1040 Contig ID leu701149079.h1 5'-most EST

Method BLASTX NCBI GI q3128175 BLAST score 176 E value 2.0e-12 Match length 99 % identity 38

(AC004521) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No. 3935

3034 2.R1040 Contig ID $1eu7\overline{0}1146445.h1$ 5'-most EST

Seq. No. 3936

3034 3.R1040 Contig ID

5'-most EST LIB3039-053-Q1-E1-F6

3937 Seq. No.

Contig ID 3035 1.R1040

5'-most EST LIB3167-029-P1-K1-D3

3938 Seq. No.

3036 1.R1040 Contig ID

5'-most EST jC-gmro02910038h11a1

Method BLASTX NCBI GI g461899 BLAST score 585 E value 4.0e-60 Match length 169 % identity 66

PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, CHLOROPLAST PRECURSOR NCBI Description

(PPIASE) (ROTAMASE) (CYCLOPHILIN) (CYCLOSPORIN A-BINDING PROTEIN) >gi 1076368 pir B53422 peptidylprolyl isomerase (EC 5.2.1.8) ROC4 - Arabidopsis thaliana >gi 405131 (L14845) cyclophilin [Arabidopsis thaliana] >gi 1322278

(U42724) cyclophilin [Arabidopsis thaliana]



```
3939
Seq. No.
                   3037 1.R1040
Contig ID
                   epx701108588.h1
5'-most EST
                  BLASTX
Method
                   g4006871
NCBI GI
BLAST score
                   700
E value
                   2.0e-73
Match length
                   355
                   47
% identity
                  (Z99707) patatin-like protein [Arabidopsis thaliana]
NCBI Description
                   3940
Seq. No.
                   3041 1.R1040
Contig ID
                   jC-qmf102220140b12d1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1483228
BLAST score
                   142
E value
                   6.0e-11
Match length
                   52
                   77
% identity
NCBI Description (X99653) MADS3 protein [Betula pendula]
Seq. No.
                   3941
                   3041 2.R1040
Contig ID
5'-most EST
                   LIB3109-022-Q1-K1-G8
Method
                   BLASTX
                   g1483228
NCBI GI
BLAST score
                   237
                   7.0e-20
E value
Match length
                   78
% identity
                   63
NCBI Description (X99653) MADS3 protein [Betula pendula]
                   3942
Seq. No.
                   3043 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220127a03a1
Seq. No.
                   3943
Contig ID
                   3047 1.R1040
5'-most EST
                   k117\overline{0}1206723.h1
Method
                   BLASTX
NCBI GI
                   g2501497
BLAST score
                   466
E value
                   1.0e-48
Match length
                   258
% identity
                   40
                   FLAVONOL 3-O-GLUCOSYLTRANSFERASE (UDP-GLUCOSE FLAVONOID
NCBI Description
                   3-O-GLUCOSYLTRANSFERASE) (ANTHOCYANIN RHAMNOSYL
                   TRANSFERASE)
                   3944
Seq. No.
Contig ID
                   3048 1.R1040
```

Contig ID 5'-most EST 3048_1.R1040 pcp700994804.h1

Seq. No.

3945

Contig ID 5'-most EST

3048_2.R1040 vzy700756210.h1



```
3946
Seq. No.
                   3049 1.R1040
Contig ID
                   g5057499
5'-most EST
                   BLASTX
Method
                   g2982469
NCBI GI
BLAST score
                   255
E value
                   8.0e-22
                   58
Match length
                   84
% identity
                  (AL022223) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3947
                   3050 1.R1040
Contig ID
                   g4293435
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3980393
                   397
BLAST score
                   2.0e-38
E value
                   130
Match length
                   56
% identity
                   (AC004561) putative glutathione S-transferase [Arabidopsis
NCBI Description
                   thaliana]
                   3948
Seq. No.
                   3051 1.R1040
Contig ID
                   crh7\overline{0}0855414.h1
5'-most EST
Seq. No.
                   3949
                   3052 1.R1040
Contig ID
                   uC-gmflminsoy065f12b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4249382
BLAST score
                   1169
                   0.0e+00
E value
Match length
                   533
                   75
% identity
NCBI Description
                   (AC005966) Strong similarity to gi 3337350 F13P17.3
                   putative permease from Arabidopsis thaliana BAC
                   gb_AC004481. [Arabidopsis thaliana]
Seq. No.
Contig ID
                   3052 2.R1040
5'-most EST
                   rry700808361.h1
                   BLASTX
Method
NCBI GI
                   q4249382
BLAST score
                   571
E value
                   5.0e-59
Match length
                   140
% identity
                   76
```

NCBI Description

(AC005966) Strong similarity to gi 3337350 F13P17.3 putative permease from Arabidopsis thaliana BAC

gb AC004481. [Arabidopsis thaliana]

3951 Seq. No.

3052 3.R1040 Contig ID 5'-most EST pmv700889172.h1



Method BLASTX
NCBI GI g4249382
BLAST score 656
E value 1.0e-68
Match length 175
% identity 77

NCBI Description (AC005966) Strong similarity to gi_3337350 F13P17.3 putative permease from Arabidopsis thaliana BAC

gb AC004481. [Arabidopsis thaliana]

Seq. No. 3952

Contig ID 3052_4.R1040 5'-most EST asn701131917.h1

Method BLASTX
NCBI GI 94249382
BLAST score 302
E value 2.0e-27
Match length 72
% identity 75

NCBI Description (AC005966) Strong similarity to gi 3337350 F13P17.3

putative permease from Arabidopsis thaliana BAC

gb_AC004481. [Arabidopsis thaliana]

Seq. No. 3953

Contig ID 3052_5.R1040 5'-most EST uC-gmropic089b11b1

Method BLASTX
NCBI GI 94249382
BLAST score 476
E value 8.0e-48
Match length 113
% identity 81

NCBI Description (AC005966) Strong similarity to gi_3337350 F13P17.3

putative permease from Arabidopsis thaliana BAC

gb AC004481. [Arabidopsis thaliana]

Seq. No. 3954

Contig ID 3052 6.R1040

5'-most EST LIB3139-089-P1-N1-F8

Method BLASTX
NCBI GI g4249382
BLAST score 283
E value 2.0e-25
Match length 70
% identity 81

NCBI Description (AC005966) Strong similarity to gi_3337350 F13P17.3

putative permease from Arabidopsis thaliana BAC

gb AC004481. [Arabidopsis thaliana]

Seq. No. 3955

Contig ID 3052 8.R1040 5'-most EST bth700849121.h1

Method BLASTN
NCBI GI g4519191
BLAST score 41
E value 7.0e-14
Match length 105



```
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:
NCBI Description
                  K9P8, complete sequence
                   3956
Seq. No.
                  3052 9.R1040
Contig ID
5'-most EST
                  eep700869859.h1
Method
                  BLASTX
                  g4249382
NCBI GI
BLAST score
                  321
                   7.0e-30
E value
                  71
Match length
% identity
                   87
                   (AC005966) Strong similarity to gi 3337350 F13P17.3
NCBI Description
                   putative permease from Arabidopsis thaliana BAC
                   gb AC004481. [Arabidopsis thaliana]
Seq. No.
                   3957
                   3052 10.R1040
Contig ID
5'-most EST
                  uC-gmropic104b03b1
                   3958
Seq. No.
                   3052 11.R1040
Contig ID
                  sat701008109.hl
5'-most EST
                   3959
Seq. No.
                   3054 1.R1040
Contig ID
5'-most EST
                  LIB3028-036-Q1-B1-A11
                   3960
Seq. No.
Contig ID
                   3056 1.R1040
                   LIB3028-028-Q1-B1-D7
5'-most EST
Method
                   BLASTX
NCBI GI
                   q1209756
BLAST score
                   694
                   4.0e-73
E value
                   162
Match length
% identity
NCBI Description
                   (U43629) integral membrane protein [Beta vulgaris]
                   3961
Seq. No.
Contig ID
                   3060 1.R1040
5'-most EST
                   LIB3092-020-Q1-K1-A9
Seq. No.
                   3962
Contia ID
                   3061 1.R1040
5'-most EST
                   leu701148824.hl
```

Method BLASTN NCBI GI g4406531 BLAST score 346 0.0e + 00E value 977 Match length % identity 85

Vigna radiata NADPH-protochlorophyllide oxidoreductase NCBI Description mRNA, chloroplast gene encoding chloroplast protein,

complete cds



Seq. No. 3963

Contig ID 3061_2.R1040
5'-most EST g5057756
Method BLASTN
NCBI GI g4406531
BLAST score 160
E value 1.0e-84
Match length 429
% identity 87

% identity 87
NCBI Description Vigna radiata NADPH-protochlorophyllide oxidoreductase

mRNA, chloroplast gene encoding chloroplast protein,

complete cds

Seq. No. 3964

Contig ID 3061 3.R1040

5'-most EST LIB3092-032-Q1-K1-B12

Method BLASTN
NCBI GI 94406531
BLAST score 156
E value 4.0e-82
Match length 431
% identity 93

NCBI Description Vigna radiata NADPH-protochlorophyllide oxidoreductase

mRNA, chloroplast gene encoding chloroplast protein,

complete cds

Seq. No. 3965

3061 4.R1040 Contig ID 5'-most EST g5752868 BLASTN Method NCBI GI g4406531 BLAST score 206 1.0e-112 E value Match length 322 % identity 91

NCBI Description Vigna radiata NADPH-protochlorophyllide oxidoreductase

mRNA, chloroplast gene encoding chloroplast protein,

complete cds

Seq. No. 3966

Contig ID 3064 1.R1040

5'-most EST uC-gmrominsoy115f07b1

Method BLASTX
NCBI GI g4538947
BLAST score 440
E value 4.0e-43
Match length 133
% identity 24

NCBI Description (ALO49483) putative mitochondrial carrier protein

[Arabidopsis thaliana]

Seq. No. 3967

Contig ID 3066_1.R1040

5'-most EST uC-gmrominsoy188c08b1

Method BLASTX NCBI GI g3253095 BLAST score 881

Match length

NCBI Description

% identity

115 52



```
9.0e-95
E value
                   307
Match length
                   57
% identity
                   (AB015643) Polygalacturonase-inhibiting protein [Citrus sp.
NCBI Description
                   cv. sannumphung]
Seq. No.
                   3968
                   3069 1.R1040
Contig ID
                  LIB3040-047-Q1-E1-C9
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4467119
BLAST score
                   1303
E value
                   1.0e-144
                   261
Match length
                   90
% identity
                   (AL035538) Histone deacetylase [Arabidopsis thaliana]
NCBI Description
                   3969
Seq. No.
                   3070 1.R1040
Contig ID
                   LIB3106-095-Q1-K1-D9
5'-most EST
                   3970
Seq. No.
                   3071 1.R1040
Contig ID
                   LIB3051-094-Q1-K1-C4
5'-most EST
Method
                   BLASTX
                   q2275196
NCBI GI
BLAST score
                   693
                   5.0e-73
E value
Match length
                   171
                   78
% identity
                   (AC002337) water stress-induced protein, WSI76 isolog
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   3971
Contig ID
                   3071 2.R1040
                   LIB3028-004-Q1-B1-E5
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2275196
BLAST score
                   150
                   4.0e-13
E value
                   67
Match length
% identity
NCBI Description
                   (AC002337) water stress-induced protein, WSI76 isolog
                   [Arabidopsis thaliana]
                   3972
Seq. No.
Contig ID
                   3072 1.R1040
5'-most EST
                   uC-gmflminsoy026a05b1
Method
                   BLASTX
NCBI GI
                   g2244993
BLAST score
                   318
                   5.0e-29
E value
```

[Arabidopsis thaliana]

(Z97341) similarity to AMP-activated protein kinase beta



```
3973
Seq. No.
                  3075 1.R1040
Contig ID
                  LIB3107-025-Q1-K1-E10
5'-most EST
Method
                  BLASTX
                  g3033397
NCBI GI
                  281
BLAST score
                   5.0e-25
E value
Match length
                  72
% identity
                   76
                   (AC004238) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3974
Seq. No.
                   3075 2.R1040
Contig ID
5'-most EST
                  LIB3106-096-Q1-K1-B2
Method
                   BLASTX
                   g3033397
NCBI GI
BLAST score
                   252
                   1.0e-21
E value
Match length
                   60
                   80
% identity
                   (AC004238) unknown protein [Arabidopsis thaliana]
NCBI Description
                   3975
Seq. No.
                   3076 1.R1040
Contig ID
5'-most EST
                   jC-gmro02910063c10a1
                   3976
Seq. No.
                   3076 2.R1040
Contig ID
5'-most EST
                   wvk700686545.hl
                   3977
Seq. No.
                   3077 1.R1040
Contig ID
                   leu701154357.h1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g3290208
BLAST score
                   126
                   3.0e-64
E value
Match length
                   230
% identity
                   89
                   Malus domestica MADS-box protein 1 mRNA, complete cds
NCBI Description
Seq. No.
Contig ID
                   3077 2.R1040
5'-most EST
                   leu701151545.h1
                   BLASTN
Method
NCBI GI
                   q3646333
BLAST score
                   91
E value
                   2.0e-43
Match length
                   189
```

90 % identity

Malus domestica mRNA for MADS box protein MdMADS8 NCBI Description

Seq. No. 3979

Contig ID 3078 1.R1040

LIB3028-035-Q1-B1-E11 5'-most EST

Method BLASTX g3068717 NCBI GI



BLAST score 1.0e-31 E value 100 Match length 65 % identity

NCBI Description (AF049236) unknown [Arabidopsis thaliana]

Seq. No.

3079 1.R1040 Contig ID

uC-gmrominsoy114c11b1 5'-most EST

3982

3983

3984

3980

Seq. No.

3981 3079 2.R1040 Contig ID

5'-most EST uC-gmropic090b10b1

Seq. No.

3081 1.R1040 Contig ID 5'-most EST ncj700975617.hl

Seq. No.

3081 2.R1040 Contig ID 5'-most EST zsq701121309.h1

Seq. No.

3081 3.R1040 Contig ID 5'-most EST gsv701044734.h1

Seq. No.

Contig ID 3083 1.R1040

5'-most EST LIB3107-060-Q1-K1-F12

3985

Method BLASTX NCBI GI q4539296 BLAST score 375 E value 6.0e-36 Match length 110

% identity 57

(AL049480) putative pathogenesis-related protein NCBI Description

[Arabidopsis thaliana]

3986 Seq. No.

3085 1.R1040 Contig ID

5'-most EST LIB3028-035-Q1-B1-E9

Seq. No. 3987

3085 2.R1040 Contig ID 5'-most EST zhf700956091.h1

Seq. No.

3988

3090 1.R1040 Contig ID

jC-gmst02400030g12a1 5'-most EST

BLASTX Method NCBI GI g3785975 BLAST score 254 2.0e-21 E value Match length 248 % identity 27

NCBI Description (AC005560) hypothetical protein [Arabidopsis thaliana]



```
3989
Seq. No.
                   3093 1.R1040
Contig ID
                  pcp700995355.h1
5'-most EST
                  BLASTX
Method
                  g2500036
NCBI GI
BLAST score
                  801
E value
                   4.0e-87
Match length
                  186
                   81
% identity
                  DIHYDROOROTASE PRECURSOR (DHOASE) >gi_2121273 (AF000146)
NCBI Description
                   dihydroorotase [Arabidopsis thaliana]
                   >gi_3292818_emb_CAA19808_ (AL031018) dihydroorotase
                   [Arabidopsis thaliana]
                   3990
Seq. No.
                   3094 1.R1040
Contig ID
5'-most EST
                   gsv701045588.hl
                  BLASTN
Method
NCBI GI
                   g1041244
BLAST score
                   268
                   1.0e-149
E value
                   657
Match length
                   87
% identity
NCBI Description A.glutinosa mRNA for enolase
                   3991
Seq. No.
                   3094 2.R1040
Contig ID
                   vwf7\overline{0}0674352.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3023713
BLAST score
                   466
E value
                   2.0e-55
Match length
                   147
% identity
                   67
                   ENOLASE (2-PHOSPHOGLYCERATE DEHYDRATASE)
NCBI Description
                   (2-PHOSPHO-D-GLYCERATE HYDRO-LYASE) (OSE1) >gi 780372
                   (U09450) enolase [Oryza sativa]
                   3992
Seq. No.
Contig ID
                   3094 3.R1040
5'-most EST
                   q429\overline{3}723
Method
                   BLASTN
NCBI GI
                   q1041244
BLAST score
                   142
E value
                   6.0e-74
Match length
                   375
% identity
NCBI Description
                   A.glutinosa mRNA for enolase
                   3993
Seq. No.
                   3094 6.R1040
Contig ID
                   kmv700741338.h1
5'-most EST
                   BLASTX
```

Method BLASTX
NCBI GI g3912997
BLAST score 147
E value 2.0e-09
Match length 37



% identity NCBI Description

FLORAL HOMEOTIC PROTEIN AGL15 >gi_2129535_pir__S71200 AGL15 protein - Arabidopsis thaliana >gi_790635 (U22528) AGL15

[Arabidopsis thaliana]

Seq. No. 3994

Contig ID 3094_7.R1040 5'-most EST txt700734472.h1

Method BLASTN
NCBI GI g533473
BLAST score 36
E value 6.0e-11
Match length 68

% identity 88

NCBI Description Mesembryanthemum crystallinum 2-phospho-D-glycerate

hydrolase, enolase, mRNA, complete cds

Seq. No. 3995

Contig ID 3095_1.R1040 5'-most EST zzp700832658.h1

Seq. No. 3996

Contig ID 3096 1.R1040

5'-most EST jC-gmf102220079b08d1

Seq. No. 3997

Contig ID 3096 2.R1040

5'-most EST jC-gmf102220080b03d1

Seq. No. 3998

Contig ID 3099 1.R1040 5'-most EST fua701043125.h1

Method BLASTX
NCBI GI g586765
BLAST score 175
E value 4.0e-12
Match length 79
% identity 48

% identity 48
NCBI Description HYPOTHETICAL 10.2 KD PROTEIN IN RPS1-NAD4L INTERGENIC

REGION (ORF 86A) >gi_282786_pir__S25987 hypothetical protein 86a - liverwort (Marchantia polymorpha)

mitochondrion >gi 786215 (M68929) ORF86a [Marchantia

polymorpha]

Seq. No. 3999

Contig ID 3100 1.R1040 5'-most EST leu701150468.h1

Seq. No. 4000

Contig ID 3100 2.R1040

5'-most EST LIB3139-113-P1-N1-B10

Seq. No. 4001

Contig ID 3100 3.R1040

5'-most EST LIB3049-012-Q1-E1-F7

Seq. No. 4002

NCBI Description



```
Contig ID
                   3102 1.R1040
5'-most EST
                   LIB3028-023-Q1-B1-E2
                   BLASTX
Method
                   g2244964
NCBI GI
BLAST score
                   743
E value
                   1.0e-78
                   233
Match length
% identity
                   61
                   (Z97340) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4003
                   3104 1.R1040
Contig ID
                   LIB3093-030-Q1-K1-E2
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4185141
BLAST score
                   1794
                   0.0e+00
E value
Match length
                   655
% identity
                   (AC005724) putative calmodulin-binding protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   4004
                   3104 2.R1040
Contig ID
                   vzy700754865.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4185141
BLAST score
                   214
                   5.0e-17
E value
Match length
                   64
% identity
NCBI Description
                   (AC005724) putative calmodulin-binding protein [Arabidopsis
                   thalianal
Seq. No.
                   4005
                   3104 5.R1040
Contig ID
5'-most EST
                   epx7\overline{0}1104351.h1
Method
                   BLASTN
                   g1698547
NCBI GI
BLAST score
                   38
                   5.0e-12
E value
Match length
                   86
% identity
                   86
                   Nicotiana tabacum calmodulin-binding protein (TCB60) mRNA,
NCBI Description
                   complete cds
                   4006
Seq. No.
Contig ID
                   3107 1.R1040
5'-most EST
                   2DC - \overline{0}1 - Q1 - B1 - D7
Method
                   BLASTX
NCBI GI
                   g418399
BLAST score
                   257
E value
                   7.0e-22
Match length
                   110
% identity
                   51
```

VACUOLAR ATP SYNTHASE SUBUNIT D (V-ATPASE D SUBUNIT)

>gi 320705 pir S30826 hypothetical protein YEL051w - yeast



(Saccharomyces cerevisiae) >gi_1042219_bbs_168845 vacuolar proton-translocating ATPase V1 subunit, V H(+) -ATPase V1 subunit=VMA8 product [Saccharomyces cerevisiae=baker's yeast, Peptide, 256 aa] >gi_603628 (U18779) Vma8p: subunit D of vacuolar H-ATPase [Saccharomyces cerevisiae]

 Seq. No.
 4007

 Contig ID
 3107_2.R1040

 5'-most EST
 2DC-01-Q1-E1-D7

Method BLASTX
NCBI GI g4567281
BLAST score 388
E value 3.0e-37
Match length 167
% identity 50

NCBI Description (AC006841) unknown protein [Arabidopsis thaliana]

Seq. No. 4009

Contig ID 3110 1.R1040 5'-most EST zhf700957271.h1

Method BLASTX
NCBI GI g3776005
BLAST score 1083
E value 0.0e+00
Match length 403
% identity 89

NCBI Description (AJ010466) RNA helicase [Arabidopsis thaliana]

Seq. No. 4010

Contig ID 3110 4.R1040

5'-most EST uC-gmropic062f10b1

Seq. No. 4011

Contig ID 3110_5.R1040 5'-most EST zsg701125976.h1

Method BLASTX
NCBI GI g3776005
BLAST score 261
E value 7.0e-23
Match length 63
% identity 83

NCBI Description (AJ010466) RNA helicase [Arabidopsis thaliana]

Seq. No. 4012

Contig ID 3112_1.R1040 5'-most EST uxk700671938.h1

Method BLASTX
NCBI GI g3549665
BLAST score 263
E value 1.0e-22
Match length 206
% identity 38

NCBI Description (AL031394) hypothetical protein [Arabidopsis thaliana]



```
4013
Seq. No.
Contig ID
                   3113_1.R1040
5'-most EST
                   LIB3050-001-Q1-E1-C2
Method
                   BLASTX
NCBI GI
                   g3075391
BLAST score
                   1544
                   1.0e-172
E value
Match length
                   579
                   55
% identity
NCBI Description
                  (AC004484) unknown protein [Arabidopsis thaliana]
Seq. No.
                   4014
Contig ID
                   3113 2.R1040
5'-most EST
                   jC-gmle01810053a05d1
                   4015
Seq. No.
                   3113 3.R1040
Contig ID
5'-most EST
                   g439<del>6</del>396
Seq. No.
                   4016
Contig ID
                   3116 1.R1040
5'-most EST
                   g5677064
                   BLASTX
Method
NCBI GI
                   g2288999
BLAST score
                   221
E value
                   5.0e-18
Match length
                   44
% identity
                   86
NCBI Description
                   (AC002335) electron transfer flavoprotein ubiquinone
                   oxidoreductase isolog [Arabidopsis thaliana]
Seq. No.
                   4017
Contig ID
                   3116 2.R1040
5'-most EST
                   dpv701100529.h1
Seq. No.
                   4018
Contig ID
                   3116 4.R1040
5'-most EST
                   LIB3028-035-Q1-B1-C11
Seq. No.
                   4019
                   3117 1.R1040
Contig ID
5'-most EST
                   uC-qmflminsoy091d09b1
Method
                   BLASTX
NCBI GI
                   g3063392
BLAST score
                   1625
E value
                   0.0e+00
Match length
                   437
                   75
% identity
NCBI Description
                   (AB012932) Ca2+/H+ exchanger [Vigna radiata]
Seq. No.
                   4020
Contig ID
                   3117 2.R1040
```

5'-most EST jC-gmle01810093c11a1

Method BLASTN NCBI GI g438216 BLAST score 511



E value 0.0e+00 Match length 997 8 identity 88

NCBI Description P.sativum mRNA for T subunit of glycine decarboxylase

multi-enzyme complex

Seq. No. 4021

Contig ID 3117 3.R1040

5'-most EST LIB3030-009-Q1-B1-C9

Method BLASTN
NCBI GI g3063391
BLAST score 228
E value 1.0e-125
Match length 754
% identity 88

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No. 4022

Contig ID 3117_4.R1040 5'-most EST epx701110201.h1

Method BLASTN
NCBI GI 9438216
BLAST score 108
E value 1.0e-53
Match length 251
% identity 86

NCBI Description P.sativum mRNA for T subunit of glycine decarboxylase

multi-enzyme complex

Seq. No. 4023

Contig ID 3117 5.R1040

5'-most EST jC-gmle01810017f08a2

Method BLASTN
NCBI GI g3063391
BLAST score 69
E value 2.0e-30
Match length 269
% identity 81

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No. 4024

Contig ID 3117_6.R1040 5'-most EST hrw701063289.h1

Method BLASTN
NCBI GI 9438253
BLAST score 103
E value 9.0e-51
Match length 263
% identity 85

NCBI Description S.tuberosum mRNA for T subunit of glycine decarboxylase

multi-enzyme complex

Seq. No. 4025

Contig ID 3117_7.R1040 5'-most EST ssr700560831.h1

Method BLASTN NCBI GI g3063391



BLAST score 157 E value 8.0e-83 Match length 333 % identity 93

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No.

Contig ID 3117 8.R1040

5'-most EST uC-gmflminsoy059h12b1

4026

Method BLASTN
NCBI GI g3063391
BLAST score 296
E value 1.0e-165
Match length 551
% identity 92

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No. 4027

Contig ID 3117_14.R1040 5'-most EST fde700875255.h1

Method BLASTN
NCBI GI g3063391
BLAST score 80
E value 4.0e-37
Match length 152
% identity 93

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No.

Contig ID 3117 15.R1040

5'-most EST uC-gmflminsoy075b09b1

4028

Method BLASTN
NCBI GI g3063391
BLAST score 107
E value 5.0e-53
Match length 330
% identity 83

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No. 4029

Contig ID 3117 17.R1040

5'-most EST uC-gmflminsoy082b03b1

Method BLASTN
NCBI GI g3063391
BLAST score 32
E value 1.0e-08
Match length 52
% identity 92

NCBI Description Vigna radiata mRNA for Ca2+/H+ exchanger, complete cds

Seq. No.

Contig ID 3119 1.R1040

5'-most EST LIB3028-035-Q1-B1-A9

4030

Seq. No. 4031

Contig ID 3119 2.R1040

5'-most EST LIB3106-104-Q1-K1-E2



4032 Seq. No.

3119 3.R1040 Contig ID 5'-most EST jex700908254.h1

Method BLASTX NCBI GI q1077162 BLAST score 153 E value 7.0e-10 Match length 111 35 % identity

NCBI Description

probable membrane protein YDL015c - yeast (Saccharomyces cerevisiae) >gi_683682_emb_CAA88344_ (Z48432) homolog of rat synaptic glycoprotein SC2 (S45663) [Saccharomyces cerevisiae] >gi 1430981 emb CAA98573 (Z74063) ORF YDL015c

[Saccharomyces cerevisiae]

4033 Seq. No.

Contig ID 3119 7.R1040 dpv701099354.h1 5'-most EST

Seq. No. 4034

Contig ID 3121 1.R1040 5'-most EST fde700871532.h1

Seq. No. 4035

Contig ID 3126 1.R1040

5'-most EST LIB3139-034-P1-N1-A11

Method BLASTX NCBI GI q1729980 1027 BLAST score 1.0e-112 E value Match length 223

% identity

NCBI Description

THAUMATIN-LIKE PROTEIN PRECURSOR >gi_2129751_pir__S71175 thaumatin-like protein - Arabidopsis thaliana >gi_536825 (L34693) thaumatin-like protein [Arabidopsis thaliana]

>gi_1094863_prf__2106421A thaumatin-like protein

[Arabidopsis thaliana]

4036 Seq. No.

3126 2.R1040 Contig ID

5'-most EST LIB3028-035-Q1-B1-A1

Method BLASTX NCBI GI q1729980 BLAST score 367 8.0e-35 E value Match length 73 % identity 85

THAUMATIN-LIKE PROTEIN PRECURSOR >gi 2129751 pir S71175 NCBI Description

thaumatin-like protein - Arabidopsis thaliana >gi_536825 (L34693) thaumatin-like protein [Arabidopsis thaliana]

>gi 1094863 prf 2106421A thaumatin-like protein

[Arabidopsis thaliana]

4037 Seq. No.

Contig ID 3127 1.R1040

5'-most EST uC-gmflminsoy018e04b1

% identity

NCBI Description



```
Method
                   BLASTX
NCBI GI
                   g4371293
BLAST score
                   150
E value
                   3.0e-09
Match length
                   280
                   24
% identity
                  (AC006260) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4038
                   3128 1.R1040
Contig ID
5'-most EST
                   LIB3093-038-Q1-K1-B4
Method
                   BLASTX
NCBI GI
                   g2827709
BLAST score
                   479
                   5.0e-48
E value
Match length
                   150
% identity
                   61
                  (AL021684) predicted protein [Arabidopsis thaliana]
NCBI Description
                   4039
Seq. No.
                   3130 1.R1040
Contig ID
5'-most EST
                   LIB3028-035-Q1-B1-A2
                   4040
Seq. No.
Contig ID
                   3132 1.R1040
5'-most EST
                   crh700850157.h1
                   4041
Seq. No.
                   3132 2.R1040
Contig ID
5'-most EST
                   ncj7\overline{0}0982157.h1
Method
                   BLASTX
NCBI GI
                   g4220481
BLAST score
                   141
E value
                   9.0e-09
Match length
                   66
% identity
                   53
NCBI Description
                  (AC006069) unknown protein [Arabidopsis thaliana]
Seq. No.
                   4042
                   3136 1.R1040
Contig ID
5'-most EST
                   leu701146141.h1
Seq. No.
                   4043
Contig ID
                   3136 2.R1040
5'-most EST
                   uC-gmflminsoy022g09b1
Seq. No.
                   4044
Contig ID
                   3137 1.R1040
5'-most EST
                   LIB3109-001-Q1-K4-C2
Method
                   BLASTX
NCBI GI
                   g2245378
BLAST score
                   1277
E value
                   1.0e-141
Match length
                   353
                  74
```

(U83245) auxin response factor 1 [Arabidopsis thaliana]



```
Seq. No.
Contig ID
                  3137 2.R1040
5'-most EST
                  xpa700794082.hl
                  BLASTX
Method
NCBI GI
                  q2245378
BLAST score
                  873
                  4.0e-94
E value
                  183
Match length
% identity
NCBI Description (U83245) auxin response factor 1 [Arabidopsis thaliana]
                  4046
Seq. No.
                  3137 3.R1040
Contig ID
                  kl1701207433.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2245378
BLAST score
                  389
E value
                  1.0e-37
                  93
Match length
                  77
% identity
NCBI Description (U83245) auxin response factor 1 [Arabidopsis thaliana]
                   4047
Seq. No.
                   3138 1.R1040
Contig ID
5'-most EST
                   jC-gmro02910037g07a1
                  BLASTX
Method
NCBI GI
                  q2739046
BLAST score
                  127
                   1.0e-14
E value
Match length
                   190
% identity
                  (AF024652) polyphosphoinositide binding protein Ssh2p
NCBI Description
                   [Glycine max]
                   4048
Seq. No.
                   3138 2.R1040
Contig ID
5'-most EST
                  bth700844103.h1
                   4049
Seq. No.
                   3138 3.R1040
Contig ID
                   vwf700675341.h1
5'-most EST
Method
                   BLASTX
                   g3687237
NCBI GI
                   655
BLAST score
                   1.0e-68
E value
Match length
                   167
% identity
                   71
NCBI Description (AC005169) putative Cys3His zinc-finger protein
                   [Arabidopsis thaliana]
Seq. No.
                   4050
                   3138 4.R1040
Contig ID
```

5'-most EST LIB3028-034-Q1-B1-G10 BLASTX Method NCBI GI q1710858 BLAST score 215

1.0e-16

E value



Match length 208 % identity 28

NCBI Description PUTATIVE SEC14 CYTOSOLIC FACTOR

(PHOSPHATIDYLINOSITOL/PHOSPHATIDYL-CHOLINE TRANSFER PROTEIN) (PI/PC TP) >gi_1177668_emb_CAA93167_ (Z69086)

unknown [Schizosaccharomyces pombe]

Seq. No. 4051

Contig ID 3145_1.R1040

5'-most EST LIB3028-034-Q1-B1-G7

Method BLASTX
NCBI GI g4490738
BLAST score 144
E value 6.0e-09
Match length 127
% identity 36

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 4052

Contig ID 3146_1.R1040 5'-most EST sat701008976.h1

Method BLASTX
NCBI GI g3033375
BLAST score 395
E value 3.0e-38
Match length 142
% identity 58

NCBI Description (AC004238) putative berberine bridge enzyme [Arabidopsis

thaliana]

Seq. No. 4053

Contig ID 3147_1.R1040

5'-most EST jC-gmf102220077g05a1

Method BLASTX
NCBI GI g3435096
BLAST score 259
E value 2.0e-22
Match length 75
% identity 67

NCBI Description (AF033587) SRZ-80 [Arabidopsis thaliana]

Seq. No. 4054

Contig ID 3147 2.R1040

5'-most EST LIB3106-026-Q1-K1-F1

Method BLASTX
NCBI GI g2582643
BLAST score 408
E value 6.0e-40
Match length 108
% identity 73

NCBI Description (AJ002377) RSZp21 protein [Arabidopsis thaliana]

Seq. No. 4055

Contig ID 3150_1.R1040 5'-most EST seb700651221.h1

Method BLASTX NCBI GI g4220481



```
BLAST score
                  6.0e-17
E value
                  79
Match length
                  53
% identity
                  (AC006069) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4056
                  3154 1.R1040
Contig ID
                  LIB3028-034-Q1-B1-E3
5'-most EST
Seq. No.
                  4057
                  3155 1.R1040
Contig ID
                  uC-gmflminsoy042b11b1
5'-most EST
                  BLASTX
Method
                  g886130
NCBI GI
BLAST score
                   915
E value
                   9.0e-99
Match length
                  282
% identity
                   61
                  (U28148) putative pectinesterase [Medicago sativa]
NCBI Description
Seq. No.
                   4058
                   3156 1.R1040
Contig ID
                  LIB3106-089-Q1-K1-A12
5'-most EST
                  BLASTX
Method
                   g3759184
NCBI GI
BLAST score
                   886
                   2.0e-95
E value
Match length
                   275
% identity
                   61
                  (AB018441) phi-1 [Nicotiana tabacum]
NCBI Description
                   4059
Seq. No.
                   3156 2.R1040
Contig ID
                   dpv701100326.h1
5'-most EST
                   BLASTX
Method
                   g3759184
NCBI GI
BLAST score
                   212
                   7.0e-17
E value
                   101
Match length
% identity
                   50
                  (AB018441) phi-1 [Nicotiana tabacum]
NCBI Description
                   4060
Seq. No.
                   3156 3.R1040
Contig ID
5'-most EST
                   gsv701048028.hl
```

4061 Seq. No.

3158_1.R1040 Contig ID

5'-most EST LIB3051-075-Q1-K1-H1

Method BLASTX g1542941 NCBI GI BLAST score 1321 1.0e-173 E value 396 Match length % identity

NCBI Description (X78116) Acetoacetyl-coenzyme A thiolase [Raphanus sativus]

Contig ID

Method

NCBI GI

5'-most EST

BLAST score



```
4062
Seq. No.
                  3158 2.R1040
Contig ID
                  LIB3109-048-Q1-K1-C1
5'-most EST
                  BLASTX
Method
                  g1542941
NCBI GI
                  444
BLAST score
                  5.0e-44
E value
                  118
Match length
                  79
% identity
NCBI Description (X78116) Acetoacetyl-coenzyme A thiolase [Raphanus sativus]
Seq. No.
                  4063
                  3163 1.R1040
Contig ID
                  sat701006382.h1
5'-most EST
                   4064
Seq. No.
                   3164 1.R1040
Contig ID
                   uC-gmrominsoy238d12b1
5'-most EST
                   BLASTX
Method
                   g2191145
NCBI GI
                   442
BLAST score
                   2.0e-43
E value
                   266
Match length
% identity
                   43
                  (AF007269) A_IG002N01.4 gene product [Arabidopsis thaliana]
NCBI Description
                   4065
Seq. No.
                   3165 1.R1040
Contig ID
                   LIB3028-034-Q1-B1-F5
5'-most EST
                   BLASTX
Method
                   q2244765
NCBI GI
                   275
BLAST score
                   3.0e-24
E value
                   72
Match length
% identity
                   (Z97335) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   4066
Seq. No.
                   3166 1.R1040
Contig ID
                   LIB3106-005-Q1-K1-D4
5'-most EST
                   BLASTX
Method
NCBI GI
                   q584825
BLAST score
                   727
                   1.0e-76
E value
Match length
                   147
 % identity
                   88
                   B2 PROTEIN >gi_322726_pir__S32124 B2 protein - carrot
NCBI Description
                   >gi 297889_emb_CAA51078_ (X72385) B2 protein [Daucus
                   carota]
                   4067
 Seq. No.
```

804

3166_2.R1040 kl1701203510.h2

BLASTX

750

q584825



```
3.0e-79
E value
Match length
                  149
                  92
% identity
                  B2 PROTEIN >gi 322726 pir S32124 B2 protein - carrot
NCBI Description
                  >gi 297889 emb CAA51078 (X72385) B2 protein [Daucus
                  carota]
                  4068
Seq. No.
                  3166 3.R1040
Contig ID
5'-most EST
                  jC-gmro02910054h09d1
                  4069
Seq. No.
Contig ID
                  3166 4.R1040
5'-most EST
                  LIB3092-034-Q1-K1-G1
                  BLASTX
Method
NCBI GI
                  q584825
BLAST score
                  196
E value
                  9.0e-15
                  75
Match length
                  81
% identity
                  B2 PROTEIN >qi 322726 pir S32124 B2 protein - carrot
NCBI Description
                  >gi 297889 emb CAA51078 (X72385) B2 protein [Daucus
                  carota]
                  4070
Seq. No.
Contig ID
                  3166 5.R1040
5'-most EST
                  ncj700980819.h1
Seq. No.
                  4071
                  3166 6.R1040
Contig ID
                  jC-gmro02910075c03d1
5'-most EST
                  4072
Seq. No.
                  3167 1.R1040
Contig ID
                  fC-gmst700888545a4
5'-most EST
                  BLASTN
Method
                  q440592
NCBI GI
BLAST score
                  513
                  0.0e+00
E value
Match length
                  785
% identity
                  82
NCBI Description V.faba mRNA (VfAGPC) for ADP-glucose pyrophosphorylase
                   4073
Seq. No.
                  3168 1.R1040
Contig ID
5'-most EST
                  ncj700981229.h1
Method
                  BLASTN
                  g2565339
NCBI GI
```

133 BLAST score E value 3.0e-68 Match length 460 % identity 88

Lupinus luteus ribosomal protein S14 (rps14) mRNA, complete NCBI Description

4074 Seq. No.

3168_2.R1040 Contig ID

NCBI GI

E value

BLAST score



```
5'-most EST
                  uC-gmronoir060h06b1
                  BLASTN
Method
                  g2565339
NCBI GI
                  237
BLAST score
                  1.0e-130
E value
Match length
                  457
% identity
                  88
                  Lupinus luteus ribosomal protein S14 (rps14) mRNA, complete
NCBI Description
                  4075
Seq. No.
                  3168 3.R1040
Contig ID
                  LIB3039-011-Q1-E1-G8
5'-most EST
Method
                  BLASTN
                  g2565339
NCBI GI
BLAST score
                  228
E value
                  1.0e-125
Match length
                  452
% identity
                  88
NCBI Description Lupinus luteus ribosomal protein S14 (rps14) mRNA, complete
                  4076
Seq. No.
                  3168 4.R1040
Contig ID
                  LIB3051-098-Q1-K1-A3
5'-most EST
Method
                  BLASTN
NCBI GI
                  q2565339
BLAST score
                  161
                   3.0e-85
E value
Match length
                   313
% identity
                  Lupinus luteus ribosomal protein S14 (rps14) mRNA, complete
NCBI Description
                   4077
Seq. No.
                   3168 5.R1040
Contig ID
5'-most EST
                  LIB3170-002-Q1-K1-G2
Method
                  BLASTN
                   g2565339
NCBI GI
BLAST score
                   163
E value
                   3.0e-86
                   399
Match length
% identity
                   85
                 Lupinus luteus ribosomal protein S14 (rps14) mRNA, complete
NCBI Description
                   cds
                   4078
Seq. No.
                   3168 8.R1040
Contig ID
                   LIB3167-002-Q1-K1-H3
5'-most EST
                   4079
Seq. No.
Contig ID
                   3169 1.R1040
5'-most EST
                   uC-gmropic0001a09b1
                   BLASTX
Method
```

q2852640

4.0e-33

355



```
Match length
% identity
                  40
NCBI Description
                  (AF007157) unknown [Homo sapiens]
Seq. No.
                  4080
Contig ID
                  3171 1.R1040
5'-most EST
                  ncj700980010.h1
Seq. No.
                  4081
Contig ID
                  3176 1.R1040
5'-most EST
                  LIB3106-051-Q1-K1-F3
Method
                  BLASTX
NCBI GI
                  q3450889
BLAST score
                  1510
E value
                  1.0e-168
Match length
                  408
% identity
NCBI Description
                  (AF083890) 19S proteosome subunit 9 [Arabidopsis thaliana]
Seq. No.
Contig ID
                  3176 2.R1040
5'-most EST
                  LIB3040-023-01-E1-C6
Method
                  BLASTX
NCBI GI
                  g3450889
BLAST score
                  314
E value
                  9.0e-29
Match length
                  94
% identity
                  70
NCBI Description
                  (AF083890) 19S proteosome subunit 9 [Arabidopsis thaliana]
Seq. No.
                  4083
Contig ID
                  3179 1.R1040
                  fua701042962.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2580582
BLAST score
                  1109
E value
                  1.0e-121
Match length
                  353
% identity
                  58
                  (AF000975) 7-0-methyltransferase [Medicago sativa]
NCBI Description
                  4084
Seq. No.
Contig ID
                  3180 1.R1040
5'-most EST
                  LIB3028-034-Q1-B1-D7
Seq. No.
                  4085
                  3182 1.R1040
Contig ID
5'-most EST
                  jC-gmf102220126f08a1
Seq. No.
                  4086
Contig ID
                  3188 1.R1040
```

5'-most EST rca701002090.h1

BLASTX Method NCBI GI q2118974 BLAST score 230 1.0e-18 E value Match length



```
% identity
                  histone H1.41 - garden pea >gi_556345 (L34578) histone H1
NCBI Description
                  [Pisum sativum]
Seq. No.
                  4087
                  3188 2.R1040
Contig ID
                  LIB3170-017-Q1-K1-A4
5'-most EST
Method
                  BLASTN
                  g3204126
NCBI GI
BLAST score
                  48
                  1.0e-17
E value
                  226
Match length
% identity
                  85
                  Cicer arietinum mRNA for histone H1
NCBI Description
Seq. No.
                  4088
                  3188 3.R1040
Contig ID
                  uC-gmrominsoy243f12b1
5'-most EST
Method
                  BLASTN
                  g3204126
NCBI GI
BLAST score
                  39
                  1.0e-12
E value
Match length
                  87
% identity
                  86
NCBI Description Cicer arietinum mRNA for histone H1
                  4089
Seq. No.
                  3189 1.R1040
Contig ID
                  uC-gmrominsoy200g09b1
5'-most EST
                  BLASTX
Method
                  g4510349
NCBI GI
                  227
BLAST score
                  1.0e-18
E value
                  89
Match length
% identity
                  60
                  (AC006921) putative bZIP transcription factor [Arabidopsis
NCBI Description
                  thaliana]
                  4090
Seq. No.
Contig ID
                  3189 2.R1040
5'-most EST
                  ssr700554626.h1
                  BLASTX
Method
                  q542200
NCBI GI
BLAST score
                  552
E value
                  2.0e-56
                  196
Match length
% identity
                  hypothetical protein - garden asparagus
NCBI Description
                  >gi 452714 emb CAA54526 (X77320) unknown [Asparagus
                  officinalis]
```

4091 Seq. No.

3189 3.R1040 Contig ID

5'-most EST LIB3109-015-Q1-K1-F3

Method BLASTX NCBI GI g629733 BLAST score 1966



E value 0.0e+00 Match length 461 % identity 81

NCBI Description sulfate adenylyltransferase (EC 2.7.7.4) met3-2 - potato >gi_2130000_pir__S67498 sulfate adenylyltransferase (EC 2.7.7.4) (clone StMet3-2) - potato >gi_479090_emb_CAA55655_

(X79053) sulfate adenylyltransferase [Solanum tuberosum]

Seq. No. 4092

Contig ID 3189 4.R1040 5'-most EST zhf700954358.h1

Method BLASTN
NCBI GI g170216
BLAST score 185
E value 1.0e-99
Match length 381
% identity 87

NCBI Description Nicotiana sylvestris (clone 6PCEP52-7) ubiquitin fusion

protein (UbiCEP52-7) mRNA, complete cds

Seq. No. 4093

Contig ID 3189_5.R1040 5'-most EST zzp700831422.h1

Method BLASTX
NCBI GI g542200
BLAST score 616
E value 4.0e-64
Match length 164
% identity 69

NCBI Description hypothetical protein - garden asparagus

>gi_452714_emb_CAA54526_ (X77320) unknown [Asparagus

officinalis]

Seq. No. 4094

Contig ID 3189 6.R1040

5'-most EST LIB3072-055-Q1-K1-F11

Method BLASTX
NCBI GI g3158376
BLAST score 404
E value 3.0e-39
Match length 118
% identity 66

NCBI Description (AF035385) unknown [Arabidopsis thaliana]

Seq. No. 4095

Contig ID 3189_7.R1040

5'-most EST jC-gmst02400031b07d1

Method BLASTX
NCBI GI g4033365
BLAST score 335
E value 5.0e-31
Match length 71
% identity 83

NCBI Description (AJ223499) ATP sulfurylase [Brassica juncea]

Seq. No. 4096

Contig ID 3189_13.R1040



```
sat701004906.h1
5'-most EST
Seq. No.
                  4097
Contig ID
                  3191 1.R1040
5'-most EST
                  jC-qmle01810087a10d1
Method
                  BLASTX
NCBI GI
                  g2245092
BLAST score
                  195
E value
                  6.0e-15
Match length
                  136
% identity
                  34
                  (Z97343) unnamed protein product [Arabidopsis thaliana]
NCBI Description
                  4098
Seq. No.
Contig ID
                  3193 1.R1040
5'-most EST
                  LIB3053-003-Q1-N1-B6
Seq. No.
                  4099
Contig ID
                  3196 1.R1040
5'-most EST
                  LIB3028-034-Q1-B1-C10
Method
                  BLASTX
NCBI GI
                  q1123105
BLAST score
                  142
                  8.0e-09
E value
Match length
                  111
% identity
                  33
                  (U42438) similar to S. cerevisiae longevity-assurance
NCBI Description
                  protein 1 (SP:P38703) [Caenorhabditis elegans]
                   4100
Seq. No.
                  3198 1.R1040
Contig ID
                  LIB3040-009-Q1-E1-E5
5'-most EST
Method
                  BLASTN
NCBI GI
                  g499692
BLAST score
                  148
                   4.0e-77
E value
                  343
Match length
% identity
                  87
NCBI Description Vicia faba cyclophilin mRNA, complete cds
Seq. No.
                   4101
                   3200 1.R1040
Contig ID
                  LIB3028-033-Q1-B1-F12
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2642157
BLAST score
                   1285
                   1.0e-142
E value
                   286
Match length
% identity
                   78
NCBI Description
                  (AC003000) ankyrin-like protein [Arabidopsis thaliana]
Seq. No.
                   4102
                   3201 1.R1040
Contig ID
                   fua701042634.h1
5'-most EST
Method
                   BLASTX
```

g2498892

434

NCBI GI BLAST score



E value 4.0e-87 Match length 303 % identity 23

NCBI Description SEC13-RELATED PROTEIN

Seq. No. 4103

Contig ID 3201_2.R1040 5'-most EST zzp700833353.h1

Method BLASTX
NCBI GI g3150415
BLAST score 466
E value 1.0e-46
Match length 104
% identity 29

NCBI Description (AC004165) sec13-related protein [Arabidopsis thaliana]

>gi_3420046 (AC004680) sec13-related protein [Arabidopsis

thaliana]

Seq. No. 4104

Contig ID 3201_3.R1040

5'-most EST LIB3051-020-Q1-E1-D9

Method BLASTX
NCBI GI g3150415
BLAST score 352
E value 2.0e-33
Match length 82
% identity 79

NCBI Description (AC004165) sec13-related protein [Arabidopsis thaliana]

>gi_3420046 (AC004680) sec13-related protein [Arabidopsis

thaliana]

Seq. No. 4105

Contig ID 3203_1.R1040 5'-most EST crh700849943.h1

Seq. No. 4106

Contig ID 3203 2.R1040 5'-most EST zvp700764113.h1

Seq. No. 4107

Contig ID 3205_1.R1040 5'-most EST eep700865877.h1

Method BLASTX
NCBI GI g2529663
BLAST score 611
E value 3.0e-63
Match length 223
% identity 52

NCBI Description (AC002535) putative lysophospholipase [Arabidopsis

thaliana] >gi_3738277 (AC005309) putative lysophospholipase

[Arabidopsis thaliana]

Seq. No. 4108

Contig ID 3208_1.R1040 5'-most EST g4292897

Seq. No. 4109



```
3210 1.R1040
Contig ID
5'-most EST
                  LIB3106-068-P1-K1-A7
                  BLASTX
Method
                  q2267567
NCBI GI
BLAST score
                  379
E value
                  2.0e-36
Match length
                  86
                  83
% identity
                  (AF009003) glycine-rich RNA binding protein 1 [Pelargonium
NCBI Description
                  x hortorum] > gi_2267569 (AF009004) glycine-rich RNA binding
                  protein 2 [Pelargonium x hortorum]
                  4110
Seq. No.
                  3210 5.R1040
Contig ID
                  ssr700553813.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2267567
BLAST score
                  226
E value
                  3.0e-29
Match length
                  84
% identity
                  (AF009003) glycine-rich RNA binding protein 1 [Pelargonium
NCBI Description
                  x hortorum] >gi_2267569 (AF009004) glycine-rich RNA binding
                  protein 2 [Pelargonium x hortorum]
                   4111
Seq. No.
                   3212 1.R1040
Contig ID
                   leu701149930.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q535454
BLAST score
                   541
                   3.0e-55
E value
Match length
                   147
                   68
% identity
                  (U13940) cysteine proteinase [Alnus glutinosa]
NCBI Description
Seq. No.
                   4112
                   3213 1.R1040
Contig ID
                   LIB3092-045-Q1-K1-F11
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1711355
BLAST score
                   405
                   1.0e-41
E value
                   208
Match length
                   46
% identity
                   SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE
NCBI Description
                   >gi_421786_pir__S34678 short-chain alcohol dehydrogenase -
```

Norway spruce >gi_395223_emb_CAA52213_ (X74115) short-chain

alcohol dehydrogenase [Picea abies]

Seq. No. 4113 Contig ID

3214 1.R1040 5'-most EST LIB3106-078-P1-K1-H7

Method BLASTX q1709825 NCBI GI

BLAST score 534 2.0e-54 E value



Match length 167
% identity 63
NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT PSAN PRECURSOR (PSI-N) >gi 1237124 (U32176) PSI-N [Arabidopsis thaliana]

Seq. No. 4114

Contig ID 3214 3.R1040

5'-most EST LIB3072-009-Q1-E1-G8

Method BLASTX
NCBI GI g1709825
BLAST score 350
E value 5.0e-33
Match length 125
% identity 62

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT PSAN PRECURSOR (PSI-N) >gi 1237124 (U32176) PSI-N [Arabidopsis thaliana]

Seq. No. 4115

Contig ID 3216_1.R1040 5'-most EST jex700906615.h1

Seq. No. 4116

3217 1.R1040 Contig ID g5606400 5'-most EST BLASTX Method g3281846 NCBI GI BLAST score 701 3.0e-73 E value Match length 623 % identity 48

NCBI Description (AJ006404) late elongated hypocotyl [Arabidopsis thaliana]

Seq. No. 4117

Contig ID 3217_2.R1040 5'-most EST rca700996583.h1

Method BLASTX
NCBI GI g3281846
BLAST score 260
E value 4.0e-22
Match length 233
% identity 31

NCBI Description (AJ006404) late elongated hypocotyl [Arabidopsis thaliana]

Seq. No. 4118

Contig ID 3217_3.R1040 5'-most EST jex700904220.h1

Seq. No. 4119

Contig ID 3217_5.R1040

5'-most EST uC-gmropic021a01b1

Seq. No. 4120

Contig ID 3217 7.R1040

5'-most EST jC-gmle01810089h08a1

Seq. No. 4121

Contig ID 3217_8.R1040



5'-most EST pmv700894463.h1 Method BLASTX

q1777443 NCBI GI BLAST score 228 E value 1.0e-18

Match length 48 % identity 83

NCBI Description (U28422) CCA1 [Arabidopsis thaliana] >gi_3510263 (AC005310)

DNA-binding protein CCA1 [Arabidopsis thaliana] >gi_4090569

(U79156) CCA1 [Arabidopsis thaliana]

Seq. No. 4122

Contig ID 3217 9.R1040 5'-most EST ncj700977410.hl

Seq. No. 4123

Contig ID 3217 12.R1040

5'-most EST LIB3138-130-Q1-N1-D2

Method BLASTN NCBI GI g3281845 BLAST score 43 .E value 6.0e-15 Match length 83 % identity 88

NCBI Description Arabidopsis thaliana mRNA for LATE ELONGATED HYPOCOTYL MYB

transcription factor

Seq. No. 4124

Contig ID 3217 16.R1040 5'-most EST zhf700953467.h1

Seq. No. 4125

Contig ID 3217 17.R1040 5'-most EST wvk700683912.h1

Seq. No. 4126

Contig ID 3219 1.R1040

5'-most EST jC-gmst02400015f08a1

Method BLASTX NCBI GI g1706082 BLAST score 836 E value 2.0e-89 Match length 313 % identity 53

NCBI Description SERINE CARBOXYPEPTIDASE II-3 PRECURSOR (CP-MII.3)

>gi_629787_pir__S44191 serine-type carboxypeptidase (EC
3.4.16.1) II-3 - barley >gi_619350_bbs_153536

CP-MII.3=serine carboxypeptidase [Hordeum vulgare=barley,

cv. Alexis, aleurone, Peptide, 516 aa]

>gi_474392_emb_CAA55478_ (X78877) serine carboxylase II-3

[Hordeum vulgare]

Seq. No.

Contig ID 3220 1.R1040

5'-most EST LIB3073-009-Q1-K1-C11 Method BLASTX

4127

NCBI GI g4544418



```
BLAST score
                  9.0e-39
E value
Match length
                  145
% identity
                  56
                  (AC006955) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4128
                  3221 1.R1040
Contig ID
                  uC-qmrominsoy103f05b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2980770
BLAST score
                  1519
E value
                  1.0e-169
Match length
                  430
% identity
                  (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4129
Contig ID
                  3221 2.R1040
5'-most EST
                  fC-qmro7000750362r1
Method
                  BLASTX
NCBI GI
                  q2980770
BLAST score
                  339
E value
                  2.0e-31
Match length
                  94
                   69
% identity
                  (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4130
                   3221 3.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy117f06b1
Method
                  BLASTX
                   q2980770
NCBI GI
                   190
BLAST score
                   3.0e-14
E value
Match length
                   77
                  56
% identity
                  (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   4131
Seq. No.
Contig ID
                   3223 1.R1040
5'-most EST
                   jC-gmf102220141g06a1
Method
                   BLASTX
NCBI GI
                   g2232354
BLAST score
                   439
E value
                   2.0e-43
Match length
                   159
                   53
% identity
                  (AF006081) UDPG glucosyltransferase [Solanum berthaultii]
NCBI Description
Seq. No.
                   4132
                   3224 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy136g03b1
```

Seq. No. 4133

Contig ID 3224_2.R1040 5'-most EST dpv701102995.h1



Seq. No. 4134

Contig ID 3225_1.R1040

5'-most EST jC-gmf102220079g05d1

Seq. No. 4135

Contig ID 3225 2.R1040

5'-most EST uC-gmflminsoy024d08b1

Seq. No. 4136

Contig ID 3226 1.R1040

5'-most EST LIB3040-047-Q1-E1-F12

Method BLASTN
NCBI GI g1666172
BLAST score 103
E value 2.0e-50
Match length 259
% identity 85

NCBI Description N.plumbaginifolia mRNA for BTF3-like transcription factor

نيد

Seq. No. 4137

Contig ID 3226_2.R1040 5'-most EST txt700736278.h1

Method BLASTX
NCBI GI g2982299
BLAST score 396
E value 2.0e-38
Match length 137
% identity 72

NCBI Description (AF051234) transcription factor BTF3 homolog [Picea

mariana]

Seq. No. 4138

Contig ID 3227_1.R1040 5'-most EST g4307584 Method BLASTX g2864616 NCBI GI BLAST score 230 E value 6.0e-19 Match length 69 % identity 65

NCBI Description (AL021811) hypothetical protein [Arabidopsis thaliana]

Seq. No. 4139

Contig ID 3228 1.R1040

5'-most EST LIB3040-019-Q1-E1-C7

Seq. No. 4140

Contig ID 3229_1.R1040

5'-most EST LIB3028-033-Q1-B1-E12

Seq. No. 4141

Contig ID 3230_1.R1040 5'-most EST zhf700955020.h1

Method BLASTX
NCBI GI g1619300
BLAST score 936



```
E value
                   1.0e-101
                   218
Match length
                   82
% identity
                  (X95269) LRR protein [Lycopersicon esculentum]
NCBI Description
Seq. No.
                   4142
Contig ID
                   3230 2.R1040
                   zzp700834740.h1
5'-most EST
                   4143
Seq. No.
                   3232 1.R1040
Contig ID
                   wrg700790475.h2
5'-most EST
                   BLASTX
Method
                   g1655637
NCBI GI
BLAST score
                   380
                   9.0e-80
E value
Match length
                   188
                   80
% identity
                  (Z54179) orf [Mus musculus]
NCBI Description
                   4144
Seq. No.
                   3232 2.R1040
Contig ID
5'-most EST
                   jsh7\overline{0}1067294.h1
                   4145
Seq. No.
                   3232 4.R1040
Contig ID
5'-most EST
                   LIB3028-015-Q1-B1-G5
Method
                   BLASTX
                   g1655637
NCBI GI
BLAST score
                   428
                   3.0e-42
E value
Match length
                   100
% identity
                  (Z54179) orf [Mus musculus]
NCBI Description
                   4146
Seq. No.
Contig ID
                   3233 1.R1040
5'-most EST
                   crh700855064.hl
Method
                   BLASTX
                   g2980770
NCBI GI
BLAST score
                   485
E value
                   1.0e-48
Match length
                   155
% identity
                   66
                  (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   4147
Seq. No.
                   3233 2.R1040
Contig ID
                   jC-gmle01810016b03d1
5'-most EST
Method
                   BLASTX
                   g2980770
NCBI GI
```

NCBI GI g2980779
BLAST score 304
E value 2.0e-27

Match length % identity

83 70

NCBI Description

(AL022198) putative protein kinase [Arabidopsis thaliana]



Seq. No. 4148 Contig ID 3235_1.R1040

5'-most EST LIB3028-032-Q1-B1-H10

Seq. No. 4149

Contig ID 3237_1.R1040 5'-most EST wrg700790785.h1

Method BLASTX
NCBI GI g3746791
BLAST score 279
E value 2.0e-24
Match length 186
% identity 32

NCBI Description (AF081788) putative spliceosome associated protein [Homo

sapiens] >gi_3985930_dbj_BAA34863_ (AB020623) DAM1 [Homo

sapiens]

Seq. No. 4150

Contig ID 3238_1.R1040

5'-most EST jC-gmf102220084a05a1

Method BLASTX
NCBI GI g3844596
BLAST score 369
E value 7.0e-35
Match length 153
% identity 47

NCBI Description (U28941) contains similarity to human copine I (GB:U83246)

[Caenorhabditis elegans]

Seq. No. 4151

Contig ID 3238 2.R1040

5'-most EST LIB3028-032-Q1-B1-H4

Method BLASTX
NCBI GI g1176658
BLAST score 399
E value 1.0e-38
Match length 158
% identity 54

NCBI Description HYPOTHETICAL 200.6 KD PROTEIN B0228.2 IN CHROMOSOME II

>gi_726363 (U23168) No definition line found

[Caenorhabditis elegans]

Seq. No. 4152

Contig ID 3238_3.R1040

5'-most EST uC-gmrominsoy108d01b1

Method BLASTX
NCBI GI g3844596
BLAST score 185
E value 1.0e-13
Match length 67
% identity 54

NCBI Description (U28941) contains similarity to human copine I (GB:U83246)

[Caenorhabditis elegans]

Seq. No. 4153

Contig ID 3238_4.R1040 5'-most EST jex700904237.h1



Seq. No. 4154

Contig ID 3239_1.R1040

5'-most EST LIB3040-022-Q1-E1-F10

Method BLASTN
NCBI GI g498895
BLAST score 120
E value 1.0e-60
Match length 392
% identity 83

NCBI Description Pisum sativum histone H2A homolog mRNA, complete cds

Seq. No. 4155

Contig ID 3239 2.R1040

5'-most EST uC-gmflminsoy058d06b1

Method BLASTN
NCBI GI g498895
BLAST score 116
E value 2.0e-58
Match length 368
% identity 83

NCBI Description Pisum sativum histone H2A homolog mRNA, complete cds

Seq. No. 4156

Contig ID 3239_3.R1040

5'-most EST LIB3106-110-Q1-K1-B8

Method BLASTN
NCBI GI 9498895
BLAST score 117
E value 9.0e-59
Match length 385
% identity 83

NCBI Description Pisum sativum histone H2A homolog mRNA, complete cds

Seq. No. 4157

Contig ID 3239_4.R1040

5'-most EST uC-gmropic062d07b1

Method BLASTN
NCBI GI g498895
BLAST score 113
E value 2.0e-56
Match length 385
% identity 82

NCBI Description Pisum sativum histone H2A homolog mRNA, complete cds

Seq. No. 4158

Contig ID 3239_5.R1040 5'-most EST zvj700605109.h2

Method BLASTN
NCBI GI 9498895
BLAST score 103
E value 1.0e-50
Match length 323
% identity 83

NCBI Description Pisum sativum histone H2A homolog mRNA, complete cds

Seq. No. 4159

BLAST score

Match length

E value

364

184

1.0e-34



```
3240 1.R1040
Contig ID
                   jC-gmf102220138h05a1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3158476
BLAST score
                  1185
E value
                   1.0e-130
Match length
                   286
                   80
% identity
NCBI Description (AF067185) aquaporin 2 [Samanea saman]
                   4160
Seq. No.
                   3240 2.R1040
Contig ID
                   g4298424
5'-most EST
Method
                   BLASTN
                   g1402832
NCBI GI
BLAST score
                   98
                   2.0e-47
E value
Match length
                   290
% identity
                   83
                  Beta vulgaris plasma membrane major intrinsic protein 1
NCBI Description
                   mRNA, complete cds
Seq. No.
                   4161
                   3240 5.R1040
Contig ID
                   q431\overline{3}924
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1657948
                   348
BLAST score
                   1.0e-32
E value
Match length
                   153
% identity
                   (U73466) MipC [Mesembryanthemum crystallinum]
NCBI Description
Seq. No.
                   4162
                   3240 9.R1040
Contig ID
                   pxt700943836.h1
5'-most EST
Seq. No.
                   4163
                   3241 1.R1040
Contig ID
5'-most EST
                   sat701012981.hl
                   BLASTX
Method
                   g3776578
NCBI GI
BLAST score
                   382
                   2.0e-36
E value
                   136
Match length
                   50
% identity
                   (AC005388) ESTs gb_F13915 and gb F13916 come from this
NCBI Description
                   gene. [Arabidopsis thaliana]
Seq. No.
                   4164
Contig ID
                   3243 1.R1040
                   uaw700665294.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3252807
```



```
% identity
                  (AC004705) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   4165
Seq. No.
Contig ID
                   3247 1.R1040
5'-most EST
                   leu701154509.hl
                   BLASTX
Method
NCBI GI
                   q3212851
BLAST score
                   236
E value
                   1.0e-19
Match length
                   74
                   62
% identity
                  (AC004005) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4166
Seq. No.
Contig ID
                   3247 2.R1040
                   leu701149937.h1
5'-most EST
                   4167
Seq. No.
Contig ID
                   3248 1.R1040
5'-most EST
                   k11701208893.h1
Method
                   BLASTX
NCBI GI
                   g2244806
BLAST score
                   590
E value
                   2.0e-60
Match length
                   299
% identity
                   43
                  (Z97336) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   4168
Seq. No.
                   3248 2.R1040
Contig ID
                   q5605815
5'-most EST
                   BLASTX
Method
                   g2244806
NCBI GI
BLAST score
                   333
                   6.0e-31
E value
                   104
Match length
                   58
% identity
                  (Z97336) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4169
                   3251 1.R1040
Contig ID
                   zpv700762784.h1
5'-most EST
                   4170
Seq. No.
Contig ID
                   3252 1.R1040
                   jex7\overline{0}0908866.h1
5'-most EST
```

BLASTN Method NCBI GI q4097582 205 BLAST score E value 1.0e-111 Match length 596 % identity

Nicotiana tabacum geranylgeranylated protein NTGP3 mRNA, NCBI Description

complete cds

Seq. No. 4171



```
Contig ID
                  3252 2.R1040
                  hyd700730610.h1
5'-most EST
                  BLASTN
Method
                  g1370199
NCBI GI
                  129
BLAST score
                  3.0e-66
E value
                  321
Match length
                  85
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAC1
                  4172
Seq. No.
                  3253 1.R1040
Contig ID
5'-most EST
                  LIB3050-004-Q1-E1-H5
Method
                  BLASTX
NCBI GI
                  g3242723
                  526
BLAST score
                  1.0e-53
E value
                  158
Match length
                  59
% identity
                  (AC003040) putative acetone-cyanohydrin lyase [Arabidopsis
NCBI Description
                  thaliana]
                  4173
Seq. No.
Contig ID
                  3254 1.R1040
5'-most EST
                  LIB3040-027-Q1-E1-G3
                  BLASTN
Method
                  g1711035
NCBI GI
                  172
BLAST score
E value
                  1.0e-91
Match length
                  388
                  86
% identity
                  Pisum sativum hydroxyproline rich glycoprotein PsHRGP1
NCBI Description
                  mRNA, partial cds
Seq. No.
                  4174
Contig ID
                  3254 2.R1040
5'-most EST
                  LIB3170-001-Q1-K1-D12
                  BLASTN
Method
NCBI GI
                  q1711035
BLAST score
                  188
E value
                  1.0e-101
Match length
                  392
% identity
NCBI Description
                  Pisum sativum hydroxyproline rich glycoprotein PsHRGP1
                  mRNA, partial cds
                   4175
Seq. No.
Contig ID
                  3254 3.R1040
                  LIB3139-056-P1-N1-D3
5'-most EST
Method
                  BLASTN
NCBI GI
                  g1711035
```

Method BLASTN
NCBI GI g1711035
BLAST score 74
E value 2.0e-33
Match length 118

% identity 91

NCBI Description Pisum sativum hydroxyproline rich glycoprotein PsHRGP1

mRNA, partial cds



```
4176
Seq. No.
                   3254 4.R1040
Contig ID
                   LIB3049-048-Q1-E1-G11
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4490705
                   305
BLAST score
                   2.0e-40
E value
                   128
Match length
% identity
                   68
                   (AL035680) ribosomal protein L14-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   4177
Seq. No.
                   3255 1.R1040
Contig ID
5'-most EST
                   vzy7\overline{0}0751859.h1
Seq. No.
                   4178
                   3256 1.R1040
Contig ID
                   ssr700555088.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3941289
BLAST score
                   292
                   3.0e-26
E value
                   117
Match length
% identity
                  (AF018093) similarity to SCAMP37 [Pisum sativum]
NCBI Description
                   4179
Seq. No.
                   3256 2.R1040
Contig ID
                   LIB3049-041-Q1-E1-F3
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2443878
BLAST score
                   164
E value
                   3.0e-11
Match length
                   92
% identity
                   47
                   (AC002294) Unknown protein [Arabidopsis thaliana]
NCBI Description
                   4180
Seq. No.
                   3260 1.R1040
Contig ID
                   jC-gmle01810004e08a1
5'-most EST
Method
                   BLASTX
                   q2160148
NCBI GI
                   791
BLAST score
                   3.0e-84
E value
                   239
Match length
% identity
                   66
                   (AC000375) EST gb H37044 comes from this gene. [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   4181
                   3260 2.R1040
Contig ID
                   ejt700606288.hl
5'-most EST
```

4182

3262_1.R1040

Seq. No.

Contig ID



```
zhf700962774.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g4455800
BLAST score
                   360
E value
                   9.0e - 34
Match length
                   145
% identity
                  (Z97343) unnamed protein product [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3264 2.R1040
Contig ID
5'-most EST
                  pcp700992416.hl
                   4184
Seq. No.
Contig ID
                   3264 4.R1040
5'-most EST
                   zhf700962567.h1
Seq. No.
                   4185
                   3269 1.R1040
Contig ID
                   LIB3138-070-P1-N1-D7
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3915601
BLAST score
                   570
                   1.0e-58
E value
Match length
                   207
% identity
                   ACTIVATOR 1 38 KD SUBUNIT (REPLICATION FACTOR C 38 KD
NCBI Description
                   SUBUNIT) (A1 38 KD SUBUNIT) (RF-C 38 KD SUBUNIT) (RFC38)
                   >gi 1498259 (L07541) replication factor C, 38-kDa subunit
                   [Homo sapiens] >gi_4506489_ref_NP_002906.1_pRFC3_
                   replication factor C (activator 1) 3 (38kD)
                   4186
Seq. No.
                   3273 1.R1040
Contig ID
                   asp7\overline{0}1131893.h1
5'-most EST
Method
                   BLASTX
                   g3831453
NCBI GI
                   549
BLAST score
                   4.0e-56
E value
                   179
Match length
% identity
                   (AC005700) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4187
Seq. No.
                   3273 2.R1040
Contig ID
5'-most EST
                   zsg701119359.hl
                   BLASTX
Method
                   g2388580
NCBI GI
BLAST score
                   211
E value
                   9.0e-17
Match length
                   67
% identity
                   57
                   (AC000098) Similar to Sequence 10 from patent 5477002
NCBI Description
```

Contig ID 3274_1.R1040

4188

Seq. No.

(gb_1253956). [Arabidopsis thaliana]



```
trc700563862.h1
5'-most EST
                  BLASTX
Method
                  g2501568
NCBI GI
                  219
BLAST score
                  1.0e-17
E value
Match length
                  72
% identity
                  57
                  HYPOTHETICAL 28.8 KD PROTEIN SLL0506
NCBI Description
                  >gi 1001342_dbj BAA10829 (D64006) hypothetical protein
                   [Synechocystis sp.]
                   4189
Seq. No.
                   3275 1.R1040
Contig ID
                  uC-gmrominsoy300b11b1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3953466
BLAST score
                   1063
E value
                   1.0e-116
                   287
Match length
                   70
% identity
                  (AC002328) F20N2.11 [Arabidopsis thaliana]
NCBI Description
                   4190
Seq. No.
                   3276 1.R1040
Contig ID
                   qsv701055179.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1498053
                   835
BLAST score
                   1.0e-89
E value
Match length
                   207
                   79
% identity
NCBI Description (U64436) ribosomal protein S8 [Zea mays]
                   4191
Seq. No.
                   3280 1.R1040
Contig ID
                   LIB3073-023-Q1-K1-H9
5'-most EST
Method
                   BLASTX
                   g3242659
NCBI GI
                   960
BLAST score
                   1.0e-104
E value
Match length
                   302
                   60
% identity
                  (AB015599) spermidine synthase [Coffea arabica]
NCBI Description
                   4192
Seq. No.
                   3281 1.R1040
Contig ID
                   LIB3106-067-P1-K1-H10
5'-most EST
                   BLASTX
Method
                   g416731
NCBI GI
BLAST score
                   311
E value
                   4.0e-28
Match length
                   167
                   38
% identity
```

NCBI Description POLLEN SPECIFIC PROTEIN C13 PRECURSOR >gi_82655_pir__JQ1107

18.3K protein precursor, pollen - maize

>gi 255569 bbs 113677 (S44171) pollen specific protein [Zea

mays=corn, Peptide, 170 aa] [Zea mays]



>gi 1588669 prf 2209273A Zm13 [Zea mays]

Seq. No. 4193

Contig ID 3281 2.R1040

5'-most EST LIB3051-088-Q1-K1-C12

Seq. No. 4194

Contig ID 3281_3.R1040

5'-most EST LIB3051-062-Q1-K1-A11

Seq. No. 4195

Contig ID 3281 4.R1040

5'-most EST LIB3051-039-Q1-K1-E9

Method BLASTX
NCBI GI g416731
BLAST score 308
E value 1.0e-27
Match length 158
% identity 39

NCBI Description POLLEN SPECIFIC PROTEIN C13 PRECURSOR >gi_82655_pir__JQ1107

18.3K protein precursor, pollen - maize

>gi_255569_bbs_113677 (S44171) pollen specific protein [Zea

mays=corn, Peptide, 170 aa] [Zea mays] >gi 1588669 prf 2209273A Zm13 [Zea mays]

Seq. No. 4196

Contig ID 3281 5.R1040 q5752512 5'-most EST BLASTX Method NCBI GI g1352316 BLAST score 654 E value 3.0e-68152 Match length % identity 85

NCBI Description DR1 PROTEIN HOMOLOG >gi_633026_dbj_BAA07288_ (D38110) Dr1

[Arabidopsis thaliana]

Seq. No. 4197

Contig ID 3281_6.R1040 5'-most EST zzp700831164.h1

Method BLASTX
NCBI GI g1352316
BLAST score 552
E value 3.0e-56
Match length 154
% identity 75

NCBI Description DR1 PROTEIN HOMOLOG >gi_633026_dbj_BAA07288_ (D38110) Dr1

[Arabidopsis thaliana]

Seq. No. 4198

Contig ID 3281_14.R1040 5'-most EST jsh701065943.h1

Method BLASTX
NCBI GI g1352316
BLAST score 138
E value 1.0e-08
Match length 30



% identity DR1 PROTEIN HOMOLOG >gi_633026_dbj_BAA07288_ (D38110) Dr1 NCBI Description [Arabidopsis thaliana] 4199 Seq. No. Contig ID 3283 1.R1040 uC-gmropic097h11b1 5'-most EST BLASTX Method g1930081 NCBI GI 1531 BLAST score 1.0e-171 E value 365 Match length % identity NCBI Description (U92878) acyl-ACP thioesterase [Garcinia mangostana] 4200 Seq. No. Contig ID 3283 2.R1040 5'-most EST LIB3028-047-Q1-B1-C11 BLASTX Method q1930081 NCBI GI 291 BLAST score 3.0e - 26E value 92 Match length 63 % identity NCBI Description (U92878) acyl-ACP thioesterase [Garcinia mangostana] Seq. No. 4201 3283 3.R1040 Contig ID LIB3053-010-Q1-N1-A8 5'-most EST BLASTX Method q1930081 NCBI GI 489 BLAST score E value 4.0e-49 143 Match length % identity (U92878) acyl-ACP thioesterase [Garcinia mangostana] NCBI Description 4202 Seq. No. 3283 4.R1040 Contig ID uC-qmflminsoy029g06b1 5'-most EST Method BLASTX q1930081 NCBI GI 888 BLAST score 1.0e-130 E value 327 Match length 75 % identity NCBI Description (U92878) acyl-ACP thioesterase [Garcinia mangostana] 4203 Seq. No. 3283 5.R1040 Contig ID LIB3170-042-Q1-J1-A11 5'-most EST BLASTN Method g4104241 NCBI GI 33 BLAST score

7.0e-09

37

97

E value

Match length

% identity



```
NCBI Description Gossypium hirsutum palmitoyl-acyl carrier protein thioesterase (FatB1) mRNA, partial cds
```

Contig ID 3283 6.R1040

5'-most EST LIB3093-040-Q1-K1-E8

Seq. No. 4205

Contig ID 3283 8.R1040

5'-most EST jC-gmro02910005b10a1

Method BLASTN
NCBI GI g4104241
BLAST score 54
E value 3.0e-21
Match length 82
% identity 91

NCBI Description Gossypium hirsutum palmitoyl-acyl carrier protein

thioesterase (FatB1) mRNA, partial cds

Seq. No. 4206

Contig ID 3283 9.R1040

5'-most EST jC-gmf102220103g09d1

Seq. No. 4207

Contig ID 3283 10.R1040

5'-most EST LIB3170-030-Q1-K1-B8

Seq. No. 4208

Contig ID 3283 11.R1040

5'-most EST uC-gmflminsoy059g08b1

Method BLASTN
NCBI GI g804947
BLAST score 39
E value 2.0e-12
Match length 43

% identity 98

NCBI Description A.thaliana mRNA for acyl-(acyl carrier protein)

thioesterase

Seq. No. 4209

Contig ID 3283_12.R1040

5'-most EST jC-gmfl02220138b10a1

Seq. No. 4210

Contig ID 3283 13.R1040

5'-most EST LIB3109-037-Q1-K1-G7

Seq. No. 4211

Contig ID 3288_1.R1040 5'-most EST leu701145370.h1

Method BLASTN
NCBI GI g169106
BLAST score 442
E value 0.0e+00
Match length 906
% identity 90

NCBI Description Pisum sativum IM30 protein mRNA, complete cds



```
4212
Seq. No.
                   3292 1.R1040
Contig ID
                   leu7\overline{0}1148679.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2062167
BLAST score
                   1341
                   1.0e-148
E value
                   321
Match length
                   78
% identity
                   (ACO01645) Proline-rich protein APG isolog [Arabidopsis
NCBI Description
                   thaliana]
                   4213
Seq. No.
                   3292 2.R1040
Contig ID
                   LIB3107-013-Q1-K1-B11
5'-most EST
                   BLASTX
Method
                   g2062167
NCBI GI
                   300
BLAST score
                   6.0e-27
E value
                   68
Match length
                   82
% identity
                   (AC001645) Proline-rich protein APG isolog [Arabidopsis
NCBI Description
                   thaliana]
                   4214
Seq. No.
                   3293 1.R1040
Contig ID
                   wrg700786729.h2
5'-most EST
Method
                   BLASTN
                   q603218
NCBI GI
                   183
BLAST score
                   3.0e-98
E value
                   457
Match length
                   85
% identity
                   Medicago sativa glucose-6-phosphate dehydrogenase mRNA,
NCBI Description
                   complete cds
                    4215
Seq. No.
                    3296 1.R1040
Contig ID
5'-most EST
                    leu701153364.h1
                    BLASTX
Method
                    g2760326
NCBI GI
                    582
BLAST score
                    6.0e-60
E value
                    261
Match length
                    48
% identity
                   (AC002130) F1N21.11 [Arabidopsis thaliana]
NCBI Description
                    4216
Seq. No.
                    3297 1.R1040
 Contig ID
                    LIB3028-032-Q1-B1-D3
 5'-most EST
                    BLASTX
 Method
                    g2239262
 NCBI GI
                    444
 BLAST score
                    7.0e-44
 E value
```

207

43

Match length % identity



```
(Y13285) pectin methylesterase-like protein [Zea mays]
NCBI Description
                  4217
Seq. No.
Contig ID
                  3298 1.R1040
5'-most EST
                  jC-gmro02910047c03d1
Seq. No.
                  4218
Contig ID
                  3303 1.R1040
5'-most EST
                  seb700653702.h1
                  BLASTX
Method
NCBI GI
                  g2130052
BLAST score
                  1922
E value
                   0.0e+00
Match length
                   459
                  77
% identity
                  xylose isomerase (EC 5.3.1.5) - barley
NCBI Description
                   >gi 1296809_emb_CAA64545_ (X95257) xylose isomerase
                   [Hordeum vulgare]
Seq. No.
                   4219
Contig ID
                   3304 1.R1040
5'-most EST
                   kl1701213675.h1
                   BLASTX
Method
NCBI GI
                   g1432056
BLAST score
                   174
E value
                   3.0e-12
Match length
                   88
% identity
                   51
                  (U56834) WRKY3 [Petroselinum crispum]
NCBI Description
                   4220
Seq. No.
                   3304 2.R1040
Contig ID
                   LIB3028-032-Q1-B1-A2
5'-most EST
                   4221
Seq. No.
                   3306 1.R1040
Contig ID
5'-most EST
                   sat701010447.h1
                   BLASTX
Method
NCBI GI
                   q1931639
                   637
BLAST score
E value
                   3.0e-66
                   310
Match length
                   41
% identity
                  (U95973) lysophospholipase isolog [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4222
                   3315 1.R1040
Contig ID
                   LIB3139-104-P1-N1-D1
5'-most EST
Method
                   BLASTX
                   q462013
NCBI GI
BLAST score
                   3250
                   0.0e + 00
E value
                   816
Match length
% identity
                   ENDOPLASMIN HOMOLOG PRECURSOR (GRP94 HOMOLOG)
NCBI Description
```

>gi_542022_pir__S39558 HSP90 homolog - Madagascar
periwinkle >gi_348696 (L14594) heat shock protein 90



[Catharanthus roseus]

```
4223
Seq. No.
                   3315 2.R1040
Contig ID
5'-most EST
                   fC-gmro7000749772r1
Method
                   BLASTX
NCBI GI
                   q462013
BLAST score
                   422
                   5.0e-42
E value
                   106
Match length
```

% identity 80
NCBI Description ENDOPLASMIN HOMOLOG PRECURSOR (GRP94 HOMOLOG)
>gi 542022 pir S39558 HSP90 homolog - Madagascar

periwinkle >gi_348696 (L14594) heat shock protein 90

[Catharanthus roseus]

```
Seq. No. 4224
```

Contig ID 3315_3.R1040 5'-most EST hrw701058023.h1 Method BLASTX

Method BLASTX
NCBI GI g462013
BLAST score 229
E value 4.0e-19
Match length 52
% identity 88

NCBI Description ENDOPLASMIN HOMOLOG PRECURSOR (GRP94 HOMOLOG)

>gi_542022_pir__S39558 HSP90 homolog - Madagascar
periwinkle >gi_348696 (L14594) heat shock protein 90

[Catharanthus roseus]

Seq. No. 4225

Contig ID 3316_1.R1040 5'-most EST jex700907321.h1

Method BLASTX
NCBI GI g4006877
BLAST score 347
E value 2.0e-32
Match length 111
% identity 37

NCBI Description (Z99707) RNA-binding like protein [Arabidopsis thaliana]

Seq. No. 4226

Contig ID 3319 1.R1040

5'-most EST jC-gmle01810060g11a1

Method BLASTX
NCBI GI g3763926
BLAST score 302
E value 7.0e-27
Match length 313
% identity 35

NCBI Description (AC004450) unknown protein [Arabidopsis thaliana]

Seq. No. 4227

Contig ID 3319_4.R1040 5'-most EST zpv700758505.h1

Method BLASTX NCBI GI g3763926



BLAST score 207
E value 2.0e-16
Match length 129
% identity 29

NCBI Description (AC004450) unknown protein [Arabidopsis thaliana]

Seq. No. 4228

Contig ID 3322_1.R1040 5'-most EST pmv700891362.h1

Seq. No. 4229

Contig ID 3323_1.R1040 5'-most EST pmv700894690.h1

Method BLASTX
NCBI GI 94567205
BLAST score 582
E value 5.0e-60
Match length 154
% identity 72

NCBI Description (AC007168) putative trehalose-6-phosphate phosphatase

[Arabidopsis thaliana]

Seq. No. 4230

Contig ID 3323 2.R1040

5'-most EST uC-gmrominsoy049f06b1

Method BLASTX
NCBI GI g4567205
BLAST score 425
E value 1.0e-41
Match length 113
% identity 72

NCBI Description (AC007168) putative trehalose-6-phosphate phosphatase

[Arabidopsis thaliana]

Seq. No. 4231

Contig ID 3323_3.R1040 5'-most EST uC-gmropic046h07b1

Method BLASTX
NCBI GI g2944180
BLAST score 148
E value 2.0e-09
Match length 109

Match length 109 % identity 39

NCBI Description (AF007779) trehalose-6-phosphate phosphatase [Arabidopsis

thaliana]

Seq. No. 4232

Contig ID 3323_4.R1040

5'-most EST uC-gmrominsoy310c07b1

Seq. No. 4233

Contig ID 3326_1.R1040

5'-most EST LIB3138-016-Q1-N1-A5

Method BLASTX
NCBI GI g4115357
BLAST score 876
E value 4.0e-94



```
Match length
                  378
% identity
                  47
                  (AC005957) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4234
                  3326 2.R1040
Contig ID
                  wrq700788991.h2
5'-most EST
                  BLASTX
Method
                  q4115357
NCBI GI
BLAST score
                  265
                   6.0e-23
E value
                   69
Match length
% identity
                  (AC005957) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4235
                   3329 1.R1040
Contig ID
                   jC-gmf102220065b08a1
5'-most EST
                   BLASTX
Method
                   g1001478
NCBI GI
BLAST score
                   691
E value
                   3.0e-72
Match length
                   368
                   39
% identity
                  (D63999) hypothetical protein [Synechocystis sp.]
NCBI Description
                   4236
Seq. No.
                   3330 1.R1040
Contig ID
5'-most EST
                   epx701104170.hl
Method
                   BLASTX
                   q2388710
NCBI GI
BLAST score
                   2187
                   0.0e+00
E value
                   500
Match length
                   79
% identity
                   (AF017150) betaine aldehyde dehydrogenase [Amaranthus
NCBI Description
                   hypochondriacus]
                   4237
Seq. No.
                   3330 2.R1040
Contig ID
                   leu7\overline{0}1154014.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g118490
                   417
BLAST score
                   7.0e-82
E value
Match length
                   236
                   67
% identity
                   BETAINE-ALDEHYDE DEHYDROGENASE PRECURSOR (BADH)
NCBI Description
                   >gi 99517_pir_ S19135 betaine-aldehyde dehydrogenase (EC
                   1.2.1.8) precursor - beet >gi_17934_emb_CAA41377_ (X58463)
                   betaine aldehyd dehydrogenase [Beta vulgaris]
```

4238

Contig ID 3330 3.R1040

5'-most EST jC-gmro02910038c11d1

Method BLASTX NCBI GI g2388710



```
BLAST score
                  820
                  8.0e-88
E value
                  197
Match length
                  74
% identity
                  (AF017150) betaine aldehyde dehydrogenase [Amaranthus
NCBI Description
                  hypochondriacus]
                  4239
Seq. No.
                  3330 4.R1040
Contig ID
                  bth700844477.h1
5'-most EST
                  BLASTX
Method
                  g1169286
NCBI GI
                  143
BLAST score
                   3.0e-09
E value
                   35
Match length
                   71
% identity
                  BETAINE-ALDEHYDE DEHYDROGENASE PRECURSOR (BADH)
NCBI Description
                   >gi_1085671_pir__S49205 betaine-aldehyde dehydrogenase (EC
                   1.2.1.8) precursor - Atriplex hortensis
                   >gi 510574_emb CAA49425 (X69770) betaine-aldehyde
                   dehydrogenase [Atriplex hortensis]
                   4240
Seq. No.
                   3331 1.R1040
Contig ID
                   uaw700665055.h1
5'-most EST
                   BLASTX
Method
                   q4539322
NCBI GI
BLAST score
                   241
                   9.0e-32
E value
                   111
Match length
% identity
                   (AL035679) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3333 1.R1040
Contig ID
                   wrg700790526.h2
5'-most EST
                   BLASTX
Method
                   g3193316
NCBI GI
BLAST score
                   642
                   3.0e-95
E value
                   284
Match length
% identity
                   65
                   (AF069299) contains similarity to nucleotide sugar
NCBI Description
                   epimerases [Arabidopsis thaliana]
                   4242
Seq. No.
                   3335 1.R1040
Contia ID
                   LIB3074-039-Q1-K1-C6
 5'-most EST
                   BLASTX
Method
                   g1653488
NCBI GI
                   173
BLAST score
```

3.0e-12 E value 79 Match length 42 % identity

(D90914) hypothetical protein [Synechocystis sp.] NCBI Description

4243 Seq. No.



Contig ID 3335_2.R1040 5'-most EST leu701146566.h1

Seq. No. 424

Contig ID 3336_1.R1040

5'-most EST LIB3028-031-Q1-B1-E5

Method BLASTX
NCBI GI 94337025
BLAST score 590
E value 8.0e-61
Match length 192
% identity 59

NCBI Description (AF123253) AIM1 protein [Arabidopsis thaliana]

Seq. No.

Contig ID 3336 2.R1040

5'-most EST jC-gmf102220054c05d1

4245

Method BLASTX
NCBI GI g4337025
BLAST score 296
E value 1.0e-26
Match length 82
% identity 71

NCBI Description (AF123253) AIM1 protein [Arabidopsis thaliana]

Seq. No. 4246

Contig ID 3340_1.R1040 5'-most EST leu701156583.h1

Method BLASTN
NCBI GI g2961299
BLAST score 266
E value 1.0e-147
Match length 474
% identity 89

NCBI Description Cicer arietinum mRNA for ribosomal protein L24

Seq. No. 4247

Contig ID 3340 2.R1040

5'-most EST LIB3073-016-Q1-K1-B3

Method BLASTN
NCBI GI g2961299
BLAST score 202
E value 1.0e-109
Match length 485
% identity 85

NCBI Description Cicer arietinum mRNA for ribosomal protein L24

Seq. No. 4248

Contig ID 3341_1.R1040 5'-most EST taw700659875.h1

Seq. No. 4249

Contig ID 3343_1.R1040 5'-most EST jsh701066525.h1

Seq. No. 4250

Contig ID 3346_1.R1040

Match length

% identity

364

38



```
LIB3051-046-Q1-K1-A6
5'-most EST
Seq. No.
                  4251
Contig ID
                  3347 1.R1040
                  LIB3073-025-Q1-K1-F2
5'-most EST
                  4252
Seq. No.
                  3347 2.R1040
Contig ID
                  uC-qmflminsoy032e12b1
5'-most EST
Seq. No.
                   4253
                  3351 1.R1040
Contig ID
                  uC-gmflminsoy077f09b1
5'-most EST
Method
                  BLASTX
NCBI GI
                   q285741
BLAST score
                   1135
                   1.0e-124
E value
                   415
Match length
                   59
% identity
                  (D14550) EDGP precursor [Daucus carota]
NCBI Description
                   4254
Seq. No.
                   3351 2.R1040
Contig ID
                   leu701151416.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3850580
                   254
BLAST score
E value
                   8.0e-22
Match length
                   102
                   54
% identity
                   (AC005278) Strong similarity to gb_D14550 extracellular
NCBI Description
                   dermal glycoprotein (EDGP) precursor from Daucus carota.
                   ESTs gb 84105 and gb AI100071 come from this gene.
                   [Arabidopsis thaliana]
                   4255
Seq. No.
                   3354 1.R1040
Contig ID
5'-most EST
                   leu701153022.h1
                   BLASTX
Method
NCBI GI
                   q3894186
BLAST score
                   861
                   2.0e-92
E value
Match length
                   260
                   63
% identity
                   (AC005662) putative embryo-abundant protein [Arabidopsis
NCBI Description
                   thaliana]
                   4256
Seq. No.
Contig ID
                   3356 1.R1040
                   uC-gmropic030a08b1
5'-most EST
Method
                   BLASTX
                   q2062169
NCBI GI
BLAST score
                   1308
E value
                   1.0e-145
```

NCBI Description (AC001645) ABC transporter (PDR5-like) isolog [Arabidopsis 836

% identity

NCBI Description

91

thaliana]



```
thaliana]
                  4257
Seq. No.
                  3359 1.R1040
Contig ID
                  LIB3106-022-Q1-K1-D3
5'-most EST
                  BLASTX
Method
                  g4567207
NCBI GI
BLAST score
                  534
                   3.0e-54
E value
                  143
Match length
                   65
% identity
                   (AC007168) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4258
                   3361 1.R1040
Contig ID
                  hrw701059490.h1
5'-most EST
                   4259
Seq. No.
                   3361 2.R1040
Contig ID
                   LIB3028-031-Q1-B1-D3
5'-most EST
Seq. No.
                   4260
                   3362 1.R1040
Contig ID
                   leu701155858.h1
5'-most EST
                   BLASTX
Method
                   q3421104
NCBI GI
                   1252
BLAST score
                   1.0e-138
E value
                   268
Match length
                   88
% identity
                   (AF043531) 20S proteasome beta subunit PBB2 [Arabidopsis
NCBI Description
                   thaliana]
                   4261
Seq. No.
                   3362 2.R1040
Contig ID
                   leu701152736.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3421104
                   887
BLAST score
                   1.0e-95
E value
                   195
Match length
% identity
                   (AF043531) 20S proteasome beta subunit PBB2 [Arabidopsis
NCBI Description
                   thaliana]
                   4262
Seq. No.
                   3362 3.R1040
Contig ID
5'-most EST
                   LIB3170-053-Q1-K1-F10
Method
                   BLASTX
                   g3421104
NCBI GI
BLAST score
                   721
                   2.0e-76
E value
Match length
                   148
```

(AF043531) 20S proteasome beta subunit PBB2 [Arabidopsis



```
4263
Seq. No.
                  3362 4.R1040
Contig ID
5'-most EST
                  LIB3065-005-Q1-N1-F8
Method
                  BLASTX
NCBI GI
                  g3421104
BLAST score
                  435
E value
                  5.0e-43
Match length
                  107
                  75
% identity
                  (AF043531) 20S proteasome beta subunit PBB2 [Arabidopsis
NCBI Description
                  thaliana]
                   4264
Seq. No.
Contig ID
                   3362 5.R1040
5'-most EST
                   jC-gmro02910037b10a1
Method
                  BLASTX
NCBI GI
                   q3421104
BLAST score
                   522
E value
                   3.0e-66
Match length
                   141
% identity
                  (AF043531) 20S proteasome beta subunit PBB2 [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   4265
Contig ID
                   3363 1.R1040
5'-most EST
                   leu701156343.hl
Seq. No.
                   4266
                   3364 1.R1040
Contig ID
                   rca701001508.hl
5'-most EST
Method
                   BLASTX
                   g2194137
NCBI GI
BLAST score
                   291
E value
                   2.0e-25
                   89
Match length
                   63
% identity
                   (AC002062) ESTs gb_R29947,gb_H76702 come from this gene.
NCBI Description
                   [Arabidopsis thaliana]
                   4267
Seq. No.
                   3364 2.R1040
Contig ID
                   LIB3107-032-Q1-K1-G12
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2194137
                   350
BLAST score
                   6.0e-33
E value
Match length
                   131
                   59
% identity
                   (AC002062) ESTs gb_R29947,gb_H76702 come from this gene.
NCBI Description
                   [Arabidopsis thaliana]
```

4268

3364 3.R1040 Contig ID 5'-most EST wrg700785911.h2

Seq. No. 4269



```
3366 1.R1040
Contig ID
                  LIB3049-006-Q1-E1-H6
5'-most EST
Seq. No.
                  3370 1.R1040
Contig ID
5'-most EST
                  LIB3170-051-Q1-K1-A11
Seq. No.
                  4271
                  3371 1.R1040
Contig ID
                  zzp700836229.hl
5'-most EST
                  BLASTX
Method
                  g1707642
NCBI GI
BLAST score
                  1420
                  1.0e-158
E value
Match length
                  571
                  48
% identity
NCBI Description (Y07748) TMK [Oryza sativa]
                  4272
Seq. No.
                  3372 1.R1040
Contig ID
                  LIB3167-031-P1-K1-H6
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4098989
BLAST score
                  833
                  3.0e-89
E value
                  215
Match length
                  71
% identity
                  (U81498) phenylalkylamine binding protein homolog
NCBI Description
                   [Arabidopsis thaliana]
                   4273
Seq. No.
Contig ID
                  3377 1.R1040
                  LIB3051-010-Q1-E1-B11
5'-most EST
Method
                  BLASTX
                  g1685005
NCBI GI
BLAST score
                   395
E value
                   4.0e-38
                   208
Match length
% identity
                   39
                  (U32644) immediate-early salicylate-induced
NCBI Description
                   qlucosyltransferase [Nicotiana tabacum]
                   4274
Seq. No.
                   3378 1.R1040
Contig ID
                   LIB3039-044-Q1-E1-E9
5'-most EST
                   BLASTX
Method
                   g416662
NCBI GI
                   216
BLAST score
                   5.0e-17
E value
```

173 Match length 40 % identity

21 KD SEED PROTEIN PRECURSOR >gi_99954_pir__S16252 trypsin inhibitor homolog - soybean >gi_21909_emb_CAA39860_ NCBI Description

(X56509) 21 kDa seed protein [Theobroma cacao]

4275 Seq. No.

3378_2.R1040 Contig ID



```
LIB3039-049-Q1-E1-D10
5'-most EST
                   4276
Seq. No.
                   3384 1.R1040
Contig ID
                   k117\overline{0}1205810.h1
5'-most EST
Seq. No.
                   4277
                   3386 1.R1040
Contig ID
5'-most EST
                   LIB3028-031-Q1-B1-C2
                   4278
Seq. No.
                   3388 1.R1040
Contig ID
                   wvk700680179.h2
5'-most EST
                   BLASTX
Method
                   g1001541
NCBI GI
                   145
BLAST score
                   4.0e-09
E value
                   63
Match length
                   46
% identity
NCBI Description (D64000) ABC transporter [Synechocystis sp.]
                   4279
Seq. No.
                   3390 1.R1040
Contig ID
                   zsq701122962.h1
5'-most EST
                   4280
Seq. No.
                   3394 1.R1040
Contig ID
                   leu7\overline{0}1156190.h1
5'-most EST
                   BLASTX
Method
                   q3170570
NCBI GI
                   155
BLAST score
                   6.0e-10
E value
Match length
                   107
                   35
% identity
NCBI Description (AF058302) FrnE [Streptomyces roseofulvus]
                   4281
Seq. No.
Contig ID
                   3394 2.R1040
                   LIB3170-058-Q1-K1-F5
5'-most EST
                   4282
Seq. No.
Contig ID
                   3398 1.R1040
                   leu701153547.h1
5'-most EST
Method
                   BLASTX
                   q3845099
NCBI GI
BLAST score
                   148
                    5.0e-09
E value
                    203
Match length
% identity
                   (AE001373) predicted secreted protein [Plasmodium
NCBI Description
                    falciparum]
Seq. No.
                    4283
```

4284

Contig ID 5'-most EST 3398 2.R1040 $gsv7\overline{0}1046010.h1$

Seq. No.



```
3399 1.R1040
Contig ID
                  uC-gmflminsoy093c05b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2664214
BLAST score
                  1307
E value
                  1.0e-144
Match length
                  633
                   47
% identity
NCBI Description (AJ222646) G2484-1 [Arabidopsis thaliana]
                   4285
Seq. No.
                  3399 5.R1040
Contig ID
                  taw700654495.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2664214
BLAST score
                   293
E value
                   4.0e-26
                   233
Match length
% identity
                   34
                  (AJ222646) G2484-1 [Arabidopsis thaliana]
NCBI Description
                   4286
Seq. No.
                   3400 1.R1040
Contig ID
                   LIB3107-022-Q1-K1-E8
5'-most EST
Method
                   BLASTN
                   g2605886
NCBI GI
BLAST score
                   75
                   1.0e-33
E value
                   225
Match length
% identity
                   87
                   Pisum sativum dormancy-associated protein (DRM1) mRNA,
NCBI Description
                   complete cds
Seq. No.
                   4287
Contia ID
                   3400 2.R1040
                   jC-gmro02910046d03a1
5'-most EST
Method
                   BLASTN
                   g2995991
NCBI GI
BLAST score
                   52
                   3.0e-20
E value
Match length
                   68
                   94
% identity
                   Arabidopsis thaliana dormancy-associated protein (DRM1)
NCBI Description
                   gene, complete cds
                   4288
Seq. No.
                   3400 4.R1040
Contig ID
                   LIB3049-013-Q1-E1-G7
5'-most EST
                   BLASTN
Method
                   g2605886
NCBI GI
BLAST score
                   56
E value
                   2.0e-22
                   172
Match length
```

Pisum sativum dormancy-associated protein (DRM1) mRNA,

88

complete cds

% identity

NCBI Description



```
4289
Seq. No.
                   3400 5.R1040
Contig ID
                  jC-qmf102220091e01a1
5'-most EST
                   4290
Seq. No.
                   3400 6.R1040
Contig ID
                   jC-qmst02400049b11a1
5'-most EST
                  BLASTN
Method
                  q2605886
NCBI GI
                   56
BLAST score
                   1.0e-22
E value
Match length
                   182
                   87
% identity
                  Pisum sativum dormancy-associated protein (DRM1) mRNA,
NCBI Description
                   complete cds
                   4291
Seq. No.
                   3401 1.R1040
Contig ID
                   LIB3109-006-Q1-K1-D12
5'-most EST
                   BLASTX
Method
                   q3914467
NCBI GI
BLAST score
                   1921
                   0.0e + 00
E value
                   482
Match length
                   79
% identity
                   26S PROTEASOME REGULATORY SUBUNIT S3 (NUCLEAR ANTIGEN 21D7)
NCBI Description
                   >gi 1864003_dbj_BAA19252_ (AB001422) 21D7 [Nicotiana
                   tabacum]
                   4292
Seq. No.
                   3401 3.R1040
Contig ID
                   uxk700669951.h1
5'-most EST
Method
                   BLASTX
                   q3914468
NCBI GI
BLAST score
                   200
                   1.0e-19
E value
Match length
                   89
% identity
                   67
                   26S PROTEASOME REGULATORY SUBUNIT S3 (NUCLEAR ANTIGEN 21D7)
NCBI Description
                   >gi_478411_pir__JQ2257 nuclear antigen 21D7 - carrot
                   >qi 217911 dbj BAA02696 (D13434) 21D7 antigen [Daucus
                   carota]
                   4293
Seq. No.
                   3403 1.R1040
Contig ID
                   leu701149929.h1
5'-most EST
                   BLASTX
Method
                   g625547
NCBI GI
                   1041
```

Method BLASTX
NCBI GI g625547
BLAST score 1041
E value 1.0e-113
Match length 246
% identity 80

NCBI Description chlorophyll a/b-binding protein type I - common tobacco

>gi_493723_emb_CAA45523_ (X64198) photosystem I

light-harvesting chlorophyll a/b-binding protein [Nicotiana

tabacum]



```
4294
Seq. No.
                  3403 2.R1040
Contig ID
                  epx7\overline{0}1104535.h1
5'-most EST
                  BLASTX
Method
                  g625547
NCBI GI
BLAST score
                  144
                   7.0e-09
E value
                   66
Match length
                   50
% identity
                  chlorophyll a/b-binding protein type I - common tobacco
NCBI Description
                   >gi 493723_emb CAA45523_ (X64198) photosystem I
                   light-harvesting chlorophyll a/b-binding protein [Nicotiana
                   tabacum]
                   4295
Seq. No.
                   3403 5.R1040
Contig ID
                   LIB3106-030-Q1-K1-G12
5'-most EST
                   BLASTX
Method
                   g625547
NCBI GI
                   301
BLAST score
                   1.0e-27
E value
                   97
Match length
                   63
% identity
                   chlorophyll a/b-binding protein type I - common tobacco
NCBI Description
                   >gi 493723 emb CAA45523 (X64198) photosystem I
                   light-harvesting chlorophyll a/b-binding protein [Nicotiana
                   tabacum]
                   4296
Seq. No.
                   3407 1.R1040
Contig ID
                   zpv700761975.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2352084
                   323
BLAST score
                   9.0e-30
E value
                   162
Match length
% identity
                   (U96613) serine/threonine kinase [Arabidopsis thaliana]
NCBI Description
                   4297
Seq. No.
                   3408 1.R1040
Contig ID
                   LIB3028-030-Q1-B1-F12
5'-most EST
                   4298
 Seq. No.
                   3409 1.R1040
 Contig ID
                   jC-gmf102220106h05a1
 5'-most EST
                   BLASTX
Method
                   g3293031
 NCBI GI
                   1371
 BLAST score
                   1.0e-152
```

NCBI Description

E value

Match length

% identity

3409 2.R1040 Contig ID

jC-gmf102220127b05d1 5'-most EST

441

55

(AJ007574) amino acid carrier [Ricinus communis]

5'-most EST

Seq. No.

Contig ID 5'-most EST



```
Method
                  BLASTX
NCBI GI
                  q4164408
BLAST score
                  218
E value
                  2.0e-17
Match length
                  66
                  50
% identity
NCBI Description
                 (AJ132228) amino acid carrier [Ricinus communis]
                  4300
Seq. No.
                  3409 3.R1040
Contig ID
5'-most EST
                  jC-qmf102220088d03a1
                  BLASTX
Method
NCBI GI
                  g4164408
BLAST score
                  149
                  1.0e-09
E value
                  46
Match length
% identity
                 (AJ132228) amino acid carrier [Ricinus communis]
NCBI Description
Seq. No.
Contig ID
                  3409 4.R1040
5'-most EST
                  pcp700991253.h1
Seq. No.
                  4302
Contig ID
                  3410 1.R1040
5'-most EST
                  uC-gmrominsoy279b04b1
Method
                  BLASTX
NCBI GI
                  q3281853
BLAST score
                  1167
                  1.0e-128
E value
Match length
                  259
                  92
% identity
                  (AL031004) putative protein [Arabidopsis thaliana]
NCBI Description
                   4303
Seq. No.
                   3410 2.R1040
Contig ID
5'-most EST
                  hyd700730125.h1
Method
                  BLASTX
NCBI GI
                   q3281853
BLAST score
                   318
                   4.0e-29
E value
Match length
                  62
% identity
                   97
                  (AL031004) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4304
Contig ID
                   3410 3.R1040
5'-most EST
                   jC-gmst02400069e07a1
Seq. No.
                   4305
Contig ID
                   3410 7.R1040
```

kmv700738573.h1

pmv700889945.hl

3410 8.R1040

4306



```
Seq. No. 4307
```

Contig ID 3411_1.R1040

5'-most EST LIB3049-004-Q1-E1-B11

Seq. No. 4308

Contig ID 3411_2.R1040 5'-most EST leu701145786.h1

Seq. No. 4309

Contig ID 3415_1.R1040 5'-most EST fua701038164.h1

Method BLASTX
NCBI GI g3281861
BLAST score 303
E value 1.0e-27
Match length 124
% identity 47

NCBI Description (AL031004) putative protein [Arabidopsis thaliana]

Seq. No. 4310

Contig ID 3415_2.R1040 5'-most EST ckk700605834.h2

Method BLASTX
NCBI GI 94572681
BLAST score 209
E value 8.0e-17
Match length 65
% identity 57

NCBI Description (AC006954) putative ubiquitin carboxyl terminal hydrolase

[Arabidopsis thaliana]

Seq. No. 4311

Contig ID 3418_1.R1040

5'-most EST uC-gmropic040a03b1

Seq. No. 4312

Contig ID 3418_3.R1040 5'-most EST rca701002429.h1

Seq. No. 4313

Contig ID 3419 1.R1040 5'-most EST leu701152577.h1

Seq. No. 4314

Contig ID 3420 1.R1040

5'-most EST LIB3051-109-Q1-K1-G4

Method BLASTX
NCBI GI g3402701
BLAST score 302
E value 9.0e-28
Match length 105
% identity 64

NCBI Description (AC004261) unknown protein [Arabidopsis thaliana]

Seq. No. 4315

Contig ID 3424_1.R1040

5'-most EST jC-gmro02910002e09a1



```
Method BLASTX
NCBI GI g4467128
BLAST score 635
E value 4.0e-66
Match length 242
% identity 49
NCBI Description (AL03553
```

NCBI Description (AL035538) putative protein [Arabidopsis thaliana]

Seq. No.

Contig ID 3431_1.R1040

5'-most EST LIB3139-004-P1-N1-B5

4316

Seq. No. 4317

Contig ID 3432 1.R1040

5'-most EST uC-gmflminsoy031b04b1

Method BLASTX
NCBI GI g2497753
BLAST score 327
E value 5.0e-30
Match length 114
% identity 54

NCBI Description NONSPECIFIC LIPID-TRANSFER PROTEIN 3 PRECURSOR (LTP 3)

>gi 1321915_emb_CAA65477_ (X96716) lipid transfer protein

[Prunus dulcis]

Seq. No. 4318

Contig ID 3433 1.R1040

5'-most EST LIB3106-006-Q1-K2-F9

Method BLASTX
NCBI GI g3660471
BLAST score 2865
E value 0.0e+00
Match length 630
% identity 85

NCBI Description (AJ001809) succinate dehydrogenase flavoprotein alpha

subunit [Arabidopsis thaliana]

Seq. No. 4319

Contig ID 3433_3.R1040

5'-most EST LIB3138-061-Q1-N1-H5

Method BLASTX
NCBI GI g3660471
BLAST score 304
E value 2.0e-27
Match length 108
% identity 56

NCBI Description (AJ001809) succinate dehydrogenase flavoprotein alpha

subunit [Arabidopsis thaliana]

Seq. No. 4320

Contig ID 3437 1.R1040

5'-most EST LIB3028-030-Q1-B1-E9

Seq. No. 4321

Contig ID 3439_1.R1040 5'-most EST trc700563801.h1

Method BLASTX



```
g2244749
NCBI GI
BLAST score
                  2170
                  0.0e + 00
E value
                  471
Match length
% identity
                  87
                  (Z97335) hydroxymethyltransferase [Arabidopsis thaliana]
NCBI Description
                  4322
Seq. No.
                  3439 2.R1040
Contig ID
5'-most EST
                  rlr700898970.h1
Method
                  BLASTX
                  g2244749
NCBI GI
BLAST score
                  264
                   2.0e-22
E value
                  146
Match length
% identity
                   83
NCBI Description (Z97335) hydroxymethyltransferase [Arabidopsis thaliana]
                   4323
Seq. No.
Contig ID
                   3439 3.R1040
5'-most EST
                   jC-qmle01810087a12a1
Method
                   BLASTX
                   g2244749
NCBI GI
BLAST score
                   611
                   1.0e-63
E value
                   128
Match length
% identity
                  (Z97335) hydroxymethyltransferase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4324
                   3441 1.R1040
Contig ID
5'-most EST
                   LIB3028-030-Q1-B1-B7
                   4325
Seq. No.
                   3442 1.R1040
Contig ID
                   fde700876469.h1
5'-most EST
Method
                   BLASTX
                   g2443329
NCBI GI
BLAST score
                   664
                   3.0e-69
E value
                   290
Match length
% identity
                   51
                  (D86122) Mei2-like protein [Arabidopsis thaliana]
NCBI Description
                   4326
Seq. No.
                   3443 1.R1040
Contig ID
5'-most EST
                   LIB3109-029-Q1-K1-F9
                   4327
Seq. No.
                   3443 2.R1040
Contig ID
                   zhf700958208.hl
5'-most EST
                   4328
Seq. No.
                   3444 1.R1040
Contig ID
                   hrw701063179.hl
5'-most EST
Method
                   BLASTX
```

g3122673

NCBI GI



```
855
BLAST score
                   6.0e-92
E value
                  204
Match length
                   79
% identity
NCBI Description 60S RIBOSOMAL PROTEIN L15 >gi_2245027_emb_CAB10447_
                   (Z97341) ribosomal protein [Arabidopsis thaliana]
Seq. No.
                   3444 2.R1040
Contig ID
                   LIB3170-004-Q1-K1-D9
5'-most EST
                   BLASTX
Method
                   q2982318
NCBI GI
                   536
BLAST score
                   2.0e-54
E value
                   182
Match length
                   61
% identity
                   (AF051244) probable 60S ribosomal protein L15 [Picea
NCBI Description
                   mariana]
                   4330
Seq. No.
                   3444 3.R1040
Contig ID
                   jC-g\overline{m}ro02910067a02a1
5'-most EST
                   BLASTX
Method
                   g3122673
NCBI GI
                   664
BLAST score
                   1.0e-85
E value
                   203
Match length
% identity
                   60S RIBOSOMAL PROTEIN L15 >gi_2245027_emb_CAB10447_
NCBI Description
                   (Z97341) ribosomal protein [Arabidopsis thaliana]
                   4331
Seq. No.
                   3444 4.R1040
Contig ID
                   LIB3106-107-Q1-K1-E7
5'-most EST
                   BLASTX
Method
                   g3122673
NCBI GI
                   830
BLAST score
E value
                   5.0e-89
                   204
Match length
                   77
 % identity
                   60S RIBOSOMAL PROTEIN L15 >gi_2245027_emb_CAB10447_
NCBI Description
                    (Z97341) ribosomal protein [Arabidopsis thaliana]
 Seq. No.
                   4332
                    3444 5.R1040
 Contig ID
                   LIB3138-070-P1-N1-D1
 5'-most EST
                   BLASTX
 Method
 NCBI GI
                   g3122673
                    267
 BLAST score
                    4.0e-23
 E value
                    55
 Match length
                    91
 % identity
                    60S RIBOSOMAL PROTEIN L15 >gi_2245027_emb_CAB10447_
 NCBI Description
```

Contig ID 3444_6.R1040

(Z97341) ribosomal protein [Arabidopsis thaliana]

Contig ID

4341

3455_1.R1040



```
vzy700753239.h1
5'-most EST
Seq. No.
                  4334
Contig ID
                  3444 9.R1040
                  leu701144830.h1
5'-most EST
Method
                  BLASTX
                  q3122673
NCBI GI
                  233
BLAST score
                   1.0e-19
E value
Match length
                   96
                   55
% identity
                   60S RIBOSOMAL PROTEIN L15 >gi_2245027_emb_CAB10447_
NCBI Description
                   (Z97341) ribosomal protein [Arabidopsis thaliana]
                   4335
Seq. No.
Contig ID
                   3446 2.R1040
                   LIB3028-030-Q1-B1-C12
5'-most EST
                   4336
Seq. No.
Contig ID
                   3447 1.R1040
5'-most EST
                   LIB3049-048-Q1-E1-A6
                   4337
Seq. No.
Contig ID
                   3447 3.R1040
5'-most EST
                   LIB3093-053-Q1-K1-E3
Seq. No.
                   4338
Contig ID
                   3448 1.R1040
5'-most EST
                   g4397081
                   4339
Seq. No.
Contig ID
                   3453 1.R1040
                   uC-gmflminsoy041f03b1
5'-most EST
                   BLASTX
Method
                   g2829204
NCBI GI
                   337
BLAST score
                   2.0e-31
E value
                   123
Match length
% identity
                   51
                   (AF044204) lipid transfer protein precursor [Gossypium
NCBI Description
                   hirsutum]
                   4340
Seq. No.
                   3454 1.R1040
Contig ID
5'-most EST
                   zhf700954310.h1
Method
                   BLASTX
                   q1362105
NCBI GI
                   529
BLAST score
E value
                   1.0e-53
Match length
                   257
% identity
                   42
                   alcohol dehydrogenase (EC 1.1.1.1) - common tobacco
NCBI Description
                   (fragment) >gi 551257 emb CAA57446 (X81853) alcohol
                   dehydrogenase [Nicotiana tabacum]
```



```
LIB3028-030-Q1-B1-D10
5'-most EST
                  BLASTX
Method
NCBI GI
                  g1708971
BLAST score
                  351
E value
                  6.0e-33
Match length
                  162
% identity
                   (R)-MANDELONITRILE LYASE ISOFORM 1 PRECURSOR
NCBI Description
                   (HYDROXYNITRILE LYASE 1) ((R)-OXYNITRILASE 1)
                  >gi_421871_pir__S32156 mandelonitrile lyase (EC 4.1.2.10) -
                  black cherry >gi_288116_emb_CAA51194_ (X72617)
                  mandelonitrile lyase [Prunus serotina] >gi_1730332 (U78814)
                   (R)-(+)-mandelonitrile lyase isoform MDL1 precursor [Prunus
                   serotina] >gi_1090776_prf__2019441A mandelonitrile lyase
                   [Prunus serotina]
Seq. No.
                   4342
Contig ID
                   3462 2.R1040
                  LIB3106-048-Q1-K1-A6
5'-most EST
                  BLASTX
Method
                   q2827712
NCBI GI
BLAST score
                   711
                   1.0e-128
E value
                   265
Match length
                   79
% identity
                   (AL021684) endoxyloglucan tranferase-like protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   4343
                   3465 1.R1040
Contig ID
                   wvk700681629.hl
5'-most EST
Method
                   BLASTX
                   q3600032
NCBI GI
                   290
BLAST score
                   5.0e-26
E value
Match length
                   161
                   42
% identity
                   (AF080119) contains similarity to tropomyosin (Pfam:
NCBI Description
                   Tropomyosin.hmm, score: 14.57) and ATP synthase (Pfam:
                   ATP-synt B.hmm, score: 10.89) [Arabidopsis thaliana]
                   4344
Seq. No.
                   3467 1.R1040
Contig ID
                   leu701154244.h1
5'-most EST
Method
                   BLASTN
                   g2961297
NCBI GI
BLAST score
                   49
                   5.0e-18
E value
                   305
Match length
% identity
NCBI Description Cicer arietinum mRNA for unidentified protein
```

Contig ID 3467_2.R1040

5'-most EST uC-gmropic011g03b1

Method BLASTN NCBI GI g2961297



```
BLAST score
                  8.0e-16
E value
                  89
Match length
                  93
% identity
NCBI Description Cicer arietinum mRNA for unidentified protein
Seq. No.
                   3467 4.R1040
Contig ID
                  uC-qmropic016g06b1
5'-most EST
                   4347
Seq. No.
                   3468 1.R1040
Contig ID
                   q431\overline{3}544
5'-most EST
                   BLASTX
Method
NCBI GI
                   g548437
BLAST score
                   164
E value
                   4.0e-11
                   103
Match length
                   40
% identity
                   OSH1 PROTEIN >qi 1078479 pir S53463 SWH1 protein (version
NCBI Description
                   1) - yeast (Saccharomyces cerevisiae) >gi_456143 (L28920)
                   Oshlp [Saccharomyces cerevisiae]
                   4348
Seq. No.
                   3469 1.R1040
Contig ID
                   LIB3030-003-Q1-B1-H5
5'-most EST
                   4349
Seq. No.
                   3472 1.R1040
Contig ID
5'-most EST
                   jC-qmro02910039g08a1
Method
                   BLASTX
                   q2245378
NCBI GI
BLAST score
                   1106
E value
                   1.0e-121
Match length
                   545
% identity
                   54
                   (U83245) auxin response factor 1 [Arabidopsis thaliana]
NCBI Description
                   4350
Seq. No.
                   3472 2.R1040
Contig ID
                   jC-gmro02910046h09d1
5'-most EST
Seq. No.
                   4351
                   3476 1.R1040
Contig ID
                   LIB3028-029-Q1-B1-G4
5'-most EST
                   BLASTX
Method
                   g3193298
NCBI GI
BLAST score
                   312
                   2.0e-28
E value
                   120
Match length
```

% identity

NCBI Description

3476 2.R1040 Contig ID 5'-most EST epx701104722.h1

(AF069298) T14P8.17 gene product [Arabidopsis thaliana]



```
4353
Seq. No.
                  3478 1.R1040
Contig ID
                  LIB3092-037-Q1-K1-D9
5'-most EST
                  BLASTX
Method
                  g4262239
NCBI GI
BLAST score
                  980
E value
                  1.0e-106
                  232
Match length
% identity
                  (AC006200) putative membrane transporter [Arabidopsis
NCBI Description
                  thaliana]
                  4354
Seq. No.
                  3480 1.R1040
Contig ID
                  LIB3170-088-Q1-K1-F7
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2245066
BLAST score
                  774
                   5.0e-82
E value
                   426
Match length
                   39
% identity
NCBI Description (Z97342) Beta-Amylase [Arabidopsis thaliana]
Seq. No.
                   3480 4.R1040
Contig ID
                   gsv701046710.hl
5'-most EST
                   4356
Seq. No.
                   3483 1.R1040
Contig ID
5'-most EST
                   euj700697977.h1
                   4357
Seq. No.
Contig ID
                   3484 1.R1040
                   sat701007584.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2499805
BLAST score
                   173
                   8.0e-12
E value
                   201
Match length
                   28
% identity
                   PROTEIN ARGININE N-METHYLTRANSFERASE 2
NCBI Description
                   >gi_1655625_emb_CAA67599_ (X99209) arginine
                   methyltransferase [Homo sapiens]
                   4358
Seq. No.
                   3484 2.R1040
Contig ID
                   LIB3051-015-Q1-E1-H10
5'-most EST
Method
                   BLASTX
                   g2499804
NCBI GI
                   388
BLAST score
                   3.0e-37
E value
                   163
Match length
% identity
                   48
```

NCBI Description

PROTEIN ARGININE N-METHYLTRANSFERASE 1 >gi_1390025 (U60882)

protein arginine N-methyltransferase [Rattus norvegicus]



```
3486 1.R1040
Contig ID
                  LIB3107-014-Q1-K1-B5
5'-most EST
                  BLASTX
Method
                  g1531758
NCBI GI
                  360
BLAST score
                  7.0e-34
E value
                  105
Match length
% identity
                  68
                  (X98772) AUX1 [Arabidopsis thaliana] >gi_3335360 (AC003028)
NCBI Description
                  unknown protein [Arabidopsis thaliana]
Seq. No.
                  3486 2.R1040
Contig ID
                  zhf700953621.h1
5'-most EST
                  4361
Seq. No.
                  3487 1.R1040
Contig ID
                  uaw700666850.h1
5'-most EST
                  BLASTX
Method
                  g2341042
NCBI GI
                  231
BLAST score
                   2.0e-18
E value
                   292
Match length
                   30
% identity
                  (AC000104) F19P19.26 [Arabidopsis thaliana]
NCBI Description
                   4362
Seq. No.
                   3488 1.R1040
Contig ID
                   smc700746571.h1
5'-most EST
                   BLASTX
Method
                   g2833220
NCBI GI
BLAST score
                   247
                   5.0e-21
E value
Match length
                   80
% identity
                   HYPOTHETICAL 35.7 KD PROTEIN IN DNL4-SLG1 INTERGENIC REGION
NCBI Description
                   >gi_2132030_pir__S61990 hypothetical protein YOR006c -
                   yeast (Saccharomyces cerevisiae) >gi_1151002 (U43491)
                   hypothetical protein UND313 [Saccharomyces cerevisiae]
                   >gi 1420098 emb CAA99194_ (Z74914) ORF YOR006c
                   [Saccharomyces cerevisiae]
Seq. No.
                   4363
Contig ID
                   3494 1.R1040
                   jC-gmst02400027g11a1
5'-most EST
                   BLASTX
Method
                   g3004549
NCBI GI
BLAST score
                   500
                   3.0e-50
E value
Match length
                   165
 % identity
                   (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4185152 (AC005724) unknown protein [Arabidopsis
```

4364

Contig ID

3494 2.R1040

thaliana]

BLAST score

Match length

E value

5.0e-78

210



```
uC-gmflminsoy055h11b1
5'-most EST
                  BLASTX
Method
                  q3004549
NCBI GI
BLAST score
                  597
                  2.0e-61
E value
                  278
Match length
% identity
                  (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                  >qi 4185152 (AC005724) unknown protein [Arabidopsis
                  thaliana]
                  4365
Seq. No.
                  3494 3.R1040
Contig ID
                  uC-gmropic019g05b1
5'-most EST
                  BLASTX
Method
                  g3004549
NCBI GI
                  187
BLAST score
                   5.0e-14
E value
                   56
Match length
% identity
                   (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4185152 (AC005724) unknown protein [Arabidopsis
                   thaliana]
                   4366
Seq. No.
                   3494 5.R1040
Contig ID
                   dpv701099759.h1
5'-most EST
                   BLASTX
Method
                   g3004549
NCBI GI
BLAST score
                   149
                   3.0e-18
E value
Match length
                   73
% identity
                   (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4185152 (AC005724) unknown protein [Arabidopsis
                   thaliana]
                   4367
Seq. No.
                   3495 1.R1040
Contig ID
                   zhf700956320.h1
5'-most EST
                   BLASTX
Method
                   q3559805
NCBI GI
BLAST score
                   473
                   3.0e-47
E value
Match length
                   97
% identity
                   (AJ006787) putative phytochelatin synthetase [Arabidopsis
NCBI Description
                   thaliana]
                   4368
Seq. No.
                   3496 1.R1040
Contig ID
                   LIB3139-113-P1-N1-E11
5'-most EST
                   BLASTX
Method
                   g2459443
NCBI GI
                   736
```



% identity

(AC002332) putative NAD(P)-dependent cholesterol NCBI Description

dehydrogenase [Arabidopsis thaliana]

Seq. No.

4369

Contig ID 5'-most EST 3496 2.R1040 epx701106813.hl

Method NCBI GI BLASTX q2459443

BLAST score E value

233 9.0e-20

Match length % identity

51 84

NCBI Description

(AC002332) putative NAD(P)-dependent cholesterol

dehydrogenase [Arabidopsis thaliana]

Seq. No.

4370

Contig ID 5'-most EST

3497 1.R1040 hrw701063210.h1

Method NCBI GI BLAST score BLASTN g3860322 110

E value Match length

1.0e-54 242

% identity NCBI Description

86 Cicer arietinum mRNA for hypothetical protein, clone

Can40-1

Seq. No.

4371

Contig ID 5'-most EST 3497 2.R1040 yza700764025.hl

Method

BLASTN

NCBI GI

q3860322

BLAST score

44

E value

3.0e-15

Match length

64

% identity

92

NCBI Description

Cicer arietinum mRNA for hypothetical protein, clone

Can40-1

Seq. No.

4372

Contig ID 5'-most EST 3498 1.R1040 g4397123

Method

BLASTX

NCBI GI BLAST score g3646451 172

E value Match length 7.0e-12

% identity

96 39

NCBI Description

(AL031603) mRNA cap methyltransferase [Schizosaccharomyces

pombe]

Seq. No. Contig ID 4373

3499 1.R1040

5'-most EST

LIB3050-008-Q1-E1-A7

Method NCBI GI

BLASTX g4079809



BLAST score 285 E value 3.0e-25 Match length 196 % identity 3

NCBI Description (AF071172) HERC2 [Homo sapiens]

Seq. No. 4374

Contig ID 3500 1.R1040

5'-most EST jC-gmf102220065e10d1

Seq. No. 4375

Contig ID 3502_1.R1040 5'-most EST ncj700981882.h1

Method BLASTX
NCBI GI g131187
BLAST score 732
E value 2.0e-77
Match length 220
% identity 67

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT III PRECURSOR (LIGHT-HARVESTING COMPLEX I 17 KD PROTEIN) (PSI-F)

>gi_72681_pir__F1SP3 photosystem I chain III precursor spinach >gi_21303 emb CAA31523_ (X13133) PSI subunit IV

preprotein (AA -77 to 154) [Spinacia oleracea]

>gi 226166_prf__1413236A photosystem I reaction center IV

[Spinacia oleracea]

Seq. No. 4376

Contig ID 3506_1.R1040

5'-most EST LIB3028-029-Q1-B1-C7

Seq. No. 4377

Contig ID 3508_1.R1040 5'-most EST zhf700965192.h1

Method BLASTX
NCBI GI g2253583
BLAST score 172
E value 3.0e-12
Match length 82
% identity 51

NCBI Description (U78721) hypothetical protein [Arabidopsis thaliana]

Seq. No. 4378

Contig ID 3514_1.R1040

5'-most EST uC-qmflminsoy053c08b1

Method BLASTX
NCBI GI g886130
BLAST score 561
E value 2.0e-80
Match length 327
% identity 47

NCBI Description (U28148) putative pectinesterase [Medicago sativa]

Seq. No. 4379

Contig ID 3514 2.R1040

5'-most EST LIB3028-029-Q1-B1-D4



Contig ID 3515_1.R1040 5'-most EST leu701151350.h1

Method BLASTN
NCBI GI g3264758
BLAST score 116
E value 2.0e-58
Match length 360
% identity 83

NCBI Description Prunus armeniaca 40S ribosomal protein S8 (RPS8) mRNA,

complete cds

Seq. No. 4381

Contig ID 3517 2.R1040

5'-most EST LIB3139-056-P1-N1-B1

Seq. No. 4382

Contig ID 3518_1.R1040

5'-most EST jC-gmro02910022g03a1

Method BLASTX
NCBI GI g3075392
BLAST score 179
E value 5.0e-16
Match length 71
% identity 66

NCBI Description (AC004484) putative steroid dehydrogenase [Arabidopsis

thaliana]

Seq. No. 4383

Contig ID 3522_1.R1040

5'-most EST LIB3056-010-Q1-N1-A1

Method BLASTX
NCBI GI g3834310
BLAST score 716
E value 1.0e-75
Match length 141
% identity 99

NCBI Description (AC005679) Similar to Ubiquitin-conjugating enzyme E2-17 KD

gb_D83004 from Homo sapiens. ESTs gb_T88233, gb_Z24464, gb_N37265, gb_H36151, gb_Z34711, gb_AA040983, and gb_T22122

come from this gene. [Arabidopsis thaliana]

Seq. No. 4384

Contig ID 3522 2.R1040

5'-most EST LIB3056-004-Q1-N1-A12

Method BLASTX
NCBI GI g3834310
BLAST score 712
E value 4.0e-75
Match length 141
% identity 99

NCBI Description (AC005679) Similar to Ubiquitin-conjugating enzyme E2-17 KD

gb_D83004 from Homo sapiens. ESTs gb_T88233, gb_Z24464, gb_N37265, gb_H36151, gb_Z34711, gb_AA040983, and gb_T22122

come from this gene. [Arabidopsis thaliana]

Seq. No. 4385

```
3522 3.R1040
Contig ID
5'-most EST
                  pxt700943960.hl
Method
                  BLASTX
NCBI GI
                  q3834310
BLAST score
                  240
                  8.0e-30
E value
Match length
                  71
                  89
% identity
                  (AC005679) Similar to Ubiquitin-conjugating enzyme E2-17 KD
NCBI Description
                  gb D83004 from Homo sapiens. ESTs gb_T88233, gb_Z24464,
                  gb N37265, gb H36151, gb Z34711, gb AA040983, and gb_T22122
                  come from this gene. [Arabidopsis thaliana]
                  4386
Seq. No.
Contig ID
                  3524 1.R1040
5'-most EST
                  LIB3028-029-Q1-B1-B12
                  4387
Seq. No.
                  3527 1.R1040
Contig ID
                  V4R-02-Q1-B1-D2
5'-most EST
Method
                  BLASTX
```

E value 7.0e-3
Match length 214
% identity 36
NCRI Description /AL031

NCBI Description (AL031824) protein transport protein sec23 homolog

[Schizosaccharomyces pombe]

 Seq. No.
 4388

 Contig ID
 3529_1.R1040

 5'-most EST
 jex700903469.h1

Seq. No. 4389

Contig ID 3530 1.R1040

5'-most EST jC-gmle01810020e10d1

Seq. No. 4390

Contig ID 3530_2.R1040

5'-most EST LIB3092-004-Q1-K1-A3

Seq. No. 4391

Contig ID 3531_1.R1040 5'-most EST crh700855207.h1

Method BLASTX
NCBI GI g3757523
BLAST score 247
E value 1.0e-20
Match length 61
% identity 74

NCBI Description (AC005167) putative transportin [Arabidopsis thaliana]

Seq. No.

Contig ID 3531 2.R1040 5'-most EST ncj700981843.h1

4392

Method BLÄSTX NCBI GI g3757523



BLAST score 345 E value 2.0e-36 Match length 104 % identity 73

NCBI Description (AC005167) putative transportin [Arabidopsis thaliana]

Seq. No. 4393

Contig ID 3533 1.R1040

5'-most EST LIB3138-119-Q1-N1-G1

Method BLASTX
NCBI GI g2598589
BLAST score 426
E value 8.0e-42
Match length 143
% identity 58

NCBI Description (Y15367) MtN19 [Medicago truncatula]

Seq. No. 4394

Contig ID 3533 2.R1040

5'-most EST LIB3051-087-Q1-K1-F1

Method BLASTX
NCBI GI g2598589
BLAST score 359
E value 4.0e-34
Match length 134
% identity 51

NCBI Description (Y15367) MtN19 [Medicago truncatula]

Seq. No. 4395

Contig ID 3533_3.R1040

5'-most EST LIB3051-062-Q1-K1-H4

Method BLASTX
NCBI GI g2598589
BLAST score 220
E value 8.0e-18
Match length 75
% identity 51

NCBI Description (Y15367) MtN19 [Medicago truncatula]

Seq. No. 4396

Contig ID 3536 1.R1040

5'-most EST LIB3051-077-Q1-K1-D11

Method BLASTN
NCBI GI g20355
BLAST score 49
E value 2.0e-18
Match length 125
% identity 85

NCBI Description Rice rgp1 mRNA for a ras-related GTP-binding protein

Seq. No. 4397

Contig ID 3537 1.R1040

5'-most EST jC-gmle01810080e09d1

Method BLASTX
NCBI GI g2739382
BLAST score 396
E value 3.0e-38



```
233
Match length
% identity
                  (AC002505) myosin heavy chain-like protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  4398
                  3538 1.R1040
Contig ID
                  uaw700660890.h1
5'-most EST
                  BLASTX
Method
                  q1256509
NCBI GI
BLAST score
                  546
                  1.0e-105
E value
                  220
Match length
                   81
% identity
NCBI Description (X92943) pectate lyase [Musa acuminata]
                   4399
Seq. No.
                   3538 2.R1040
Contig ID
                  LIB3028-029-Q1-B1-A9
5'-most EST
                  BLASTX
Method
                   g2463509
NCBI GI
                   975
BLAST score
                   1.0e-106
E value
                   226
Match length
                   77
% identity
NCBI Description (Y09541) pectate lyase [Zinnia elegans]
                   4400
Seq. No.
                   3539 1.R1040
Contig ID
5'-most EST
                   pxt700945309.h1
Method
                   BLASTX
                   q2880043
NCBI GI
BLAST score
                   942
                   1.0e-102
E value
                   272
Match length
                   65
% identity
                   (AC002340) putative 3-hydroxyisobutyryl-coenzyme A
NCBI Description
                   hydrolase [Arabidopsis thaliana]
                   4401
Seq. No.
                   3539 2.R1040
Contig ID
                   bth700843808.h1
5'-most EST
                   4402
Seq. No.
                   3542 1.R1040
Contig ID
                   zsg701125459.h1
5'-most EST
                   4403
Seq. No.
                   3543 1.R1040
Contig ID
                   LIB3028-028-Q1-B1-H8
5'-most EST
                   BLASTX
Method
                   g4263781
NCBI GI
BLAST score
                   346
                   2.0e-32
E value
                   191
Match length
                   41
% identity
NCBI Description (AC006068) putative membrane transport protein [Arabidopsis
```



thaliana]

Seq. No. 4404

Contig ID 3545_1.R1040 5'-most EST bth700843502.h1

Method BLASTX
NCBI GI g100535
BLAST score 313
E value 2.0e-28
Match length 79
% identity 78

NCBI Description hypothetical protein - swollen duckweed

>qi 1929057 emb CAA32236 (X14075) longest ORF (1) [Lemna

gibba]

Seq. No. 4405

Contig ID 3546_1.R1040

5'-most EST LIB3051-065-Q1-K1-E9

Method BLASTX
NCBI GI g481762
BLAST score 249
E value 3.0e-21
Match length 135
% identity 35

NCBI Description beta-adaptin 1 - fruit fly (Drosophila melanogaster)

>gi 434902_emb_CAA53509_ (X75910) beta-adaptin Drosophila 1

[Drosophila melanogaster]

Seq. No. 4406

Contig ID 3547 1.R1040

5'-most EST jC-gmst02400046a06a1

Method BLASTX
NCBI GI g2339978
BLAST score 1270
E value 1.0e-150
Match length 357
% identity 75

NCBI Description (Y11336) RGA1 protein [Arabidopsis thaliana]

Seq. No. 4407

Contig ID 3547_2.R1040 5'-most EST vzy700751203.h1

Method BLASTX
NCBI GI g2569938
BLAST score 217
E value 3.0e-17
Match length 47
% identity 85

NCBI Description (Y15193) GAI [Arabidopsis thaliana]

Seq. No. 4408

Contig ID 3549_1.R1040 5'-most EST trc700563015.h1

Seq. No. 4409

Contig ID 3552 1.R1040

5'-most EST LIB3107-080-Q1-K1-D10



```
BLASTX
Method
                  g4544369
NCBI GI
                   643
BLAST score
                   6.0e-67
E value
                   274
Match length
                   48
% identity
NCBI Description (AC006920) hypothetical protein [Arabidopsis thaliana]
                   4410
Seq. No.
                   3555 1.R1040
Contig ID
                   LIB3106-079-P1-K1-D3
5'-most EST
                   BLASTX
Method
                   g537313
NCBI GI
                   651
BLAST score
                   3.0e-68
E value
                   151
Match length
                   83
% identity
NCBI Description (L36159) unknown protein [Medicago sativa]
                   4411
Seq. No.
                   3558 1.R1040
Contig ID
                   leu701150138.h1
5'-most EST
                   BLASTX
Method
                   g2266666
NCBI GI
                   686
BLAST score
                   5.0e-72
E value
                   227
Match length
                   58
% identity
NCBI Description (Y14202) hypothetical protein [Hordeum vulgare]
                   4412
Seq. No.
                   3558 2.R1040
Contig ID
                   jC-g\overline{m}le01810085f08a1
5'-most EST
                   BLASTX
Method
                   q4335745
NCBI GI
                   576
BLAST score
E value
                   6.0e-59
                   308
Match length
% identity
                    (AC006284) putative hydrolase (contains an
NCBI Description
                    esterase/lipase/thioesterase active site serine domain
                    (prosite: PS50187) [Arabidopsis thaliana]
Seq. No.
                    4413
                    3558 3.R1040
Contig ID
                    uC-gmrominsoy138d07b1
5'-most EST
Method
                    BLASTX
                    g4115368
NCBI GI
                    510
BLAST score
                    1.0e-51
E value
                    170
Match length
                    55
 % identity
                   (AC005957) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                    4414
 Seq. No.
```

3558 4.R1040

LIB3170-085-Q1-J1-D6

Contig ID

5'-most EST



```
4415
Seq. No.
                  3558 5.R1040
Contig ID
                  jC-gmro02910032f01a1
5'-most EST
Method
                  BLASTX
                  g4335745
NCBI GI
BLAST score
                  221
                  2.0e-25
E value
                  162
Match length
                   41
% identity
                   (AC006284) putative hydrolase (contains an
NCBI Description
                   esterase/lipase/thioesterase active site serine domain
                   (prosite: PS50187) [Arabidopsis thaliana]
                   4416
Seq. No.
Contig ID
                   3558 8.R1040
5'-most EST
                   leu701149426.h1
Method
                   BLASTX
                   g4115367
NCBI GI
BLAST score
                   167
E value
                   1.0e-11
Match length
                   82
% identity
                   (AC005957) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4417
Seq. No.
                   3561 1.R1040
Contig ID
                   uC-gmropic108f06b1
5'-most EST
Method
                   BLASTX
                   g1149595
NCBI GI
                   207
BLAST score
                   3.0e-16
E value
                   141
Match length
                   35
% identity
                   (Z49860) 1-acyl-sn-glycerol-3-phosphate acyltransferase
NCBI Description
                   [Brassica napus]
                   4418
Seq. No.
                   3562 1.R1040
Contig ID
                   LIB3028-028-Q1-B1-G5
5'-most EST
                   BLASTX
Method
                   g4140691
NCBI GI
                   350
BLAST score
                   6.0e-33
E value
                   148
Match length
                   53
% identity
                   (AF101972) zeatin O-glucosyltransferase [Phaseolus lunatus]
NCBI Description
                   4419
Seq. No.
                   3564 1.R1040
Contig ID
                   zsg701123081.h1
5'-most EST
                   BLASTX
Method
                   g3298540
NCBI GI
                   985
BLAST score
                   1.0e-107
E value
Match length
                   239
```

79

% identity



NCBI Description (AC004681) unknown protein [Arabidopsis thaliana]

Seq. No. 4420

Contig ID 3565_2.R1040

5'-most EST LIB3040-044-Q1-E1-F8

Method BLASTX
NCBI GI g3482979
BLAST score 571
E value 2.0e-58
Match length 215
% identity 53

NCBI Description (AL031369) putative protein [Arabidopsis thaliana]

>gi_4567258_gb_AAD23672.1_AC007070_21 (AC007070)

hypothetical protein [Arabidopsis thaliana]

Seq. No. 4421

Contig ID 3565_3.R1040

5'-most EST LIB3028-028-Q1-B1-G8

Seq. No. 4422

Contig ID 3568 1.R1040

5'-most EST LIB3053-009-Q1-N1-C1

Method BLASTX
NCBI GI g2864612
BLAST score 518
E value 3.0e-52
Match length 229
% identity 48

NCBI Description (AL021811) L-ascorbate peroxidase - like protein

[Arabidopsis thaliana] >gi_4049334_emb_CAA22559_ (AL034567) L-ascorbate peroxidase-like protein [Arabidopsis thaliana]

Seq. No. 4423

Contig ID 3568 2.R1040

5'-most EST LIB3106-106-Q1-K1-E12

Seq. No. 4424

Contig ID 3571 1.R1040

5'-most EST LIB3028-028-Q1-B1-E3

Method BLASTX
NCBI GI g3894168
BLAST score 243
E value 3.0e-20
Match length 148
% identity 36

NCBI Description (AC005312) similar to phloem-specific lectin [Arabidopsis

thaliana]

Seq. No. 4425

Contig ID 3573_1.R1040

5'-most EST LIB3028-028-Q1-B1-E5

Seq. No. 4426

Contig ID 3573_2.R1040 5'-most EST jex700903945.h1

Seq. No. 4427



Contig ID 3575_1.R1040

5'-most EST LIB3049-011-Q1-E1-F7

Method BLASTX
NCBI GI g2289003
BLAST score 1619
E value 0.0e+00
Match length 496
% identity 65

NCBI Description (AC002335) membrane transporter D1 isolog [Arabidopsis

thaliana]

Seq. No. 4428

Contig ID 3575 2.R1040

5'-most EST jC-gmle01810086b11d1

Seq. No. 4429

Contig ID 3579_1.R1040

5'-most EST LIB3167-077-P1-K2-E5

Method BLASTX
NCBI GI g2582665
BLAST score 1377
E value 1.0e-153
Match length 336
% identity 82

NCBI Description (Z82983) thi [Citrus sinensis]

Seq. No. 4430

Contig ID 3579_2.R1040 5'-most EST ncj700979158.h1

Method BLASTX
NCBI GI g3212869
BLAST score 704
E value 2.0e-74
Match length 160
% identity 82

NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]

Seq. No. 4431

Contig ID 3579_3.R1040

5'-most EST jC-gmst02400052h01d1

Method BLASTN
NCBI GI g1289203
BLAST score 93
E value 1.0e-44
Match length 213
% identity 86

NCBI Description A.glutinosa mRNA for thiazole biosynthetic enzyme

Seq. No. 4432

Contig ID 3579 4.R1040

5'-most EST jC-gmro02910034h10d1

Method BLASTX
NCBI GI g3212869
BLAST score 2167
E value 0.0e+00
Match length 487
% identity 82



```
NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]
                  4433
Seq. No.
                  3579 5.R1040
Contig ID
                  LIB3106-114-Q1-K1-B9
5'-most EST
Method
                  BLASTN
NCBI GI
                  g596077
                  63
BLAST score
                  8.0e-27
E value
                  159
Match length
                  86
% identity
                  Zea mays thiamine biosynthetic enzyme (thi1-1) mRNA,
NCBI Description
                  complete cds
                  4434
Seq. No.
                  3579 7.R1040
Contig ID
5'-most EST
                  LIB3106-077-P1-K1-E6
                  BLASTX
Method
                  g2582665
NCBI GI
                  216
BLAST score
                   5.0e-17
E value
                  115
Match length
% identity
                   65
                  (Z82983) thi [Citrus sinensis]
NCBI Description
                   4435
Seq. No.
                   3579 11.R1040
Contig ID
                   jC-qmst02400043c02a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3212869
                   334
BLAST score
                   3.0e-31
E value
Match length
                   73
                   81
% identity
                  (AC004005) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4436
                   3581 1.R1040
Contig ID
                   epx701109079.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q1617270
                   2744
BLAST score
                   0.0e + 00
E value
                   659
Match length
                   76
% identity
                  (X94624) acyl-CoA synthetase [Brassica napus]
NCBI Description
                   4437
Seq. No.
                   3582 1.R1040
Contig ID
                   LIB3170-026-Q1-K1-B2
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2511594
                   1026
BLAST score
                   1.0e-112
E value
                   233
Match length
                   85
% identity
NCBI Description (Y13694) multicatalytic endopeptidase complex, proteasome
```



precursor, beta subunit [Arabidopsis thaliana] >gi_2827525_emb_CAA16533_ (AL021633) multicatalytic endopeptidase complex, proteasome precursor, beta subunit [Arabidopsis thaliana] >gi_3421099 (AF043529) 20S proteasome subunit PBA1 [Arabidopsis thaliana]

Seq. No. 4438

Contig ID 3585_1.R1040 5'-most EST zzp700830591.h1

Method BLASTX
NCBI GI g2501460
BLAST score 442
E value 4.0e-43
Match length 556
% identity 28

NCBI Description PROBABLE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE HAUSP

(UBIQUITIN THIOLESTERASE HAUSP) (UBIQUITIN-SPECIFIC

PROCESSING PROTEASE HAUSP) (DEUBIQUITINATING ENZYME HAUSP)

(HERPESVIRUS ASSOCIATED UBIQUITIN-SPECIFIC PROTEASE)

>gi_1545952_emb_CAA96580_ (Z72499) herpesvirus associated

ubiquitin-specific protease (HAUSP) [Homo sapiens]

>gi_4507857_ref_NP_003461.1_pUSP7_ Herpes virus-associated

ubiquitin-specific protease

Seq. No. 4439

Contig ID 3586_1.R1040

5'-most EST LIB3139-107-P1-N1-B1

Method BLASTX
NCBI GI g1403522
BLAST score 908
E value 5.0e-98
Match length 200
% identity 81

NCBI Description (X57187) chitinase [Phaseolus vulgaris]

Seq. No. 4440

Contig ID 3590_1.R1040 5'-most EST kl1701205931.h1

Seq. No. 4441

Contig ID 3591_1.R1040 5'-most EST vzy700756151.h1

Method BLASTX
NCBI GI g1905910
BLAST score 714
E value 1.0e-75
Match length 182
% identity 73

NCBI Description (AD000092) putative human phenylalanine tRNA synthetase

[Homo sapiens]

Seq. No. 4442

Contig ID 3593_1.R1040 5'-most EST rca701000704.h1

Method BLASTX
NCBI GI g1589913
BLAST score 596

NCBI GI BLAST score

398



```
9.0e-63
E value
                   182
Match length
                   71
% identity
                  (U69694) ATP-sulfurylase precursor [Brassica oleracea]
NCBI Description
                   4443
Seq. No.
                   3594 1.R1040
Contig ID
                   LIB3170-061-Q1-J1-B8
5'-most EST
                   BLASTX
Method
                   g3183217
NCBI GI
                   462
BLAST score
                   1.0e-45
E value
                   250
Match length
                   40
% identity
NCBI Description HYPOTHETICAL PROTEIN KIAA0103 >gi 285943 dbj BAA03493
                   (D14659) KIAA0103 [Homo sapiens]
Seq. No.
                   4444
                   3594 2.R1040
Contig ID
                   ncj7\overline{0}0987648.h1
5'-most EST
                   BLASTX
Method
                   g3319457
NCBI GI
BLAST score
                   141
                   1.0e-08
E value
                   77
Match length
                   39
% identity
                  (AF077542) contains similarity to O-linked GlcNAc
NCBI Description
                   transferases [Caenorhabditis elegans]
                   4445
Seq. No.
                   3596 1.R1040
Contig ID
                   q5677719
5'-most EST
Method
                   BLASTX
                   q4567207
NCBI GI
BLAST score
                   564
                   9.0e-58
E value
                   154
Match length
% identity
                   67
                   (AC007168) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4446
                   3596 3.R1040
Contig ID
5'-most EST
                   leu701150756.h1
Method
                   BLASTX
                   g4567207
NCBI GI
                   201
BLAST score
                   9.0e-16
E value
Match length
                   71
% identity
                   58
                   (AC007168) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4447
Seq. No.
                   3597 1.R1040
Contig ID
                   LIB3028-028-Q1-B1-D4
5'-most EST
                   BLASTX
Method
                   q3150412
```

868

Match length

% identity

160 36



```
9.0e-39
E value
Match length
                  115
% identity
                  (AC004165) putative Fe(II) transport protein [Arabidopsis
NCBI Description
                  thaliana] >gi 3420044 (AC004680) putative Fe(II) transport
                  protein [Arabidopsis thaliana]
                  4448
Seq. No.
                  3598 1.R1040
Contig ID
                  leu701150063.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4335734
BLAST score
                  458
                  1.0e-45
E value
Match length
                  151
                  41
% identity
NCBI Description (AC006248) putative calmodulin [Arabidopsis thaliana]
Seq. No.
                  4449
                  3602 1.R1040
Contig ID
                  LIB3093-006-Q1-K1-F8
5'-most EST
                  BLASTX
Method
                  q2497538
NCBI GI
                  220
BLAST score
                  1.0e-17
E value
                  92
Match length
                  55
% identity
                  PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi 466350 (L08632)
NCBI Description
                  pyruvate kinase [Glycine max]
                  4450
Seq. No.
Contig ID
                  3602 2.R1040
5'-most EST
                  LIB3051-013-Q1-E1-C8
Method
                  BLASTX
                  g2497543
NCBI GI
BLAST score
                  204
E value
                   6.0e-16
                  70
Match length
% identity
                   61
                  PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi 542061_pir__S41379
NCBI Description
                  pyruvate kinase - common tobacco >gi 444023_emb_CAA82628_
                   (Z29492) pyruvate kinase [Nicotiana tabacum]
                   4451
Seq. No.
                   3604 1.R1040
Contig ID
5'-most EST
                  epx701105730.h1
                   4452
Seq. No.
                   3608 1.R1040
Contig ID
5'-most EST
                   uC-gmropic113g01b1
Method
                   BLASTX
                   g135861
NCBI GI
BLAST score
                   137
E value
                   1.0e-10
```

NCBI Description INTERFERON-RELATED PROTEIN PC4 (TPA INDUCED SEQUENCE 7)



(TIS7 PROTEIN) >gi_321269_pir__A44989 interferon-related protein TIS7 - mouse >gi_54806_emb_CAA35258_ (X17400) TIS7 protein (AA 1-449) [Mus sp.]

Seq. No. 4453

Contig ID 3608 2.R1040

5'-most EST LIB3028-010-Q1-B1-F3

Seq. No. 4454

Contig ID 3610_1.R1040 5'-most EST zsg701118611.h1

Method BLASTX
NCBI GI g119640
BLAST score 548
E value 1.0e-104
Match length 354
% identity 55

NCBI Description 1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG (PROTEIN

E8) >gi_82109_pir__S01642 ripening protein E8 - tomato >gi_19199_emb_CAA31789_ (X13437) E8 protein [Lycopersicon

esculentum]

Seq. No. 4455

Contig ID 3614_1.R1040 5'-most EST g4397270

Seq. No. 4456

Contig ID 3617_1.R1040

5'-most EST LIB3028-028-Q1-B1-C2

Method BLASTX
NCBI GI g4406777
BLAST score 272
E value 3.0e-24
Match length 89
% identity 63

NCBI Description (AC006532) putative zinc-finger protein [Arabidopsis

thaliana]

Seq. No. 4457

3619 1.R1040 Contig ID g4395762 5'-most EST BLASTX Method g1362078 NCBI GI BLAST score 465 3.0e-46 E value 128 Match length % identity 65

NCBI Description endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG1)

- common nasturtium >gi_311835_emb_CAA48324_ (X68254)

cellulase [Tropaeolum majus]

Seq. No. 4458

Contig ID 3621_1.R1040

5'-most EST LIB3040-036-Q1-E1-D1

Method BLASTX
NCBI GI g2497753
BLAST score 318



```
5.0e-29
E value
                  115
Match length
% identity
                  NONSPECIFIC LIPID-TRANSFER PROTEIN 3 PRECURSOR (LTP 3)
NCBI Description
                  >gi 1321915_emb_CAA65477_ (X96716) lipid transfer protein
                  [Prunus dulcis]
                  4459
Seq. No.
                  3621 2.R1040
Contig ID
                  uC-qmflminsoy081b04b1
5'-most EST
Method
                  BLASTX
                  q2497753
NCBI GI
                  285
BLAST score
                   3.0e-25
E value
Match length
                   91
                   59
% identity
                  NONSPECIFIC LIPID-TRANSFER PROTEIN 3 PRECURSOR (LTP 3)
NCBI Description
                   >gi 1321915 emb CAA65477_ (X96716) lipid transfer protein
                   [Prunus dulcis]
                   4460
Seq. No.
Contig ID
                   3624 1.R1040
                  LIB3139-021-P1-N1-A8
5'-most EST
                   4461
Seq. No.
                   3627 1.R1040
Contig ID
5'-most EST
                  LIB3049-030-Q1-E1-H4
                  BLASTX
Method
NCBI GI
                   q4572675
BLAST score
                   154
E value
                   1.0e-09
Match length
                   79
% identity
                   47
                   (AC006954) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4462
Contig ID
                   3627 2.R1040
                   LIB3028-028-Q1-B1-A11
5'-most EST
Method
                   BLASTX
                   g4572675
NCBI GI
BLAST score
                   154
                   5.0e-10
E value
Match length
                   79
                   47
% identity
                   (AC006954) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4463
Seq. No.
                   3628 1.R1040
Contig ID
                   leu701148293.h1
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g1762933
BLAST score 265
E value 1.0e-22
Match length 188
% identity 39

NCBI Description (U66263) tumor-related protein [Nicotiana tabacum]

```
Seq. No.
                   3628 2.R1040
Contig ID
5'-most EST
                  bth700845465.h1
Seq. No.
Contig ID
                   3628 3.R1040
5'-most EST
                  bth700849368.h1
                   4466
Seq. No.
                   3629 1.R1040
Contig ID
5'-most EST
                   LIB3028-028-Q1-B1-A3
                   4467
Seq. No.
                   3631 1.R1040
Contig ID
                   ncj7\overline{0}0987071.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3850587
BLAST score
                   509
                   3.0e-51
E value
                   221
Match length
% identity
                   (AC005278) Strong similarity to gi 2244780 hypothetical
NCBI Description
                   protein from Arabidopsis thaliana chromosome 4 contig
                   gb_Z97335. [Arabidopsis thaliana]
                   4468
Seq. No.
Contig ID
                   3631 2.R1040
5'-most EST
                   kl1701214264.h1
Method
                   BLASTX
                   g2244779
NCBI GI
BLAST score
                   373
                   1.0e-35
E value
```

130 Match length 58 % identity

(Z97335) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 4469 3631 3.R1040 Contig ID LIB3028-028-Q1-B1-A4 5'-most EST Method BLASTX g2244780 NCBI GI 174 BLAST score

2.0e-12 E value Match length 67 54 % identity

(Z97335) hypothetical protein [Arabidopsis thaliana] NCBI Description

4470 Seq. No. 3635_1.R1040 Contig ID 5'-most EST g5678054 Method BLASTX q3108053 NCBI GI 1536 BLAST score 1.0e-171 E value

Match length 324 92 % identity

NCBI Description (AF056326) myo-inositol 1-phosphate synthase; INO1 [Zea



mays]

4471 Seq. No. Contig ID 3635 2.R1040 5'-most EST g4397429 Method BLASTN q602564 NCBI GI BLAST score 316 1.0e-177 E value 987 Match length % identity 83 NCBI Description C.paradisi (Macf) INO1 gene 4472

Seq. No.

Contig ID

3635 3.R1040

5'-most EST

LIB3055-012-Q1-N1-H8

Method BLASTN NCBI GI g602564 BLAST score 102 5.0e-50 E value 310 Match length % identity 83

NCBI Description C.paradisi (Macf) INO1 gene

Seq. No.

4473

Contig ID

3635 4.R1040

5'-most EST

jC-gmst02400014d06a1

BLASTN Method NCBI GI q602564 BLAST score 167 E value 8.0e-89 Match length 431 % identity 85

NCBI Description C.paradisi (Macf) INO1 gene

Seq. No.

4474

Contig ID

3635 5.R1040

5'-most EST

 $uC\hbox{-} gmflminsoy 059b 08b 1$

BLASTX Method NCBI GI q1170567 BLAST score 595 1.0e-61 E value Match length 124 89 % identity

NCBI Description

MYO-INOSITOL-1-PHOSPHATE SYNTHASE (IPS)

>gi 1085960 pir S52648 INO1 protein - Citrus paradisi >gi 602565 emb CAA83565 (Z32632) INO1 [Citrus x paradisi]

Seq. No.

4475

Contig ID 3635 6.R1040

5'-most EST

LIB3107-036-Q1-K1-G12

BLASTN Method NCBI GI q602564 BLAST score 105 7.0e-52 E value Match length 282 % identity 84



```
NCBI Description C.paradisi (Macf) INO1 gene
Seq. No.
                  4476
Contig ID
                  3635 8.R1040
5'-most EST
                  pmv700889442.h1
Method
                  BLASTX
NCBI GI
                  q1170567
BLAST score
                  193
E value
                  2.0e-23
Match length
                  62
% identity
                  97
                  MYO-INOSITOL-1-PHOSPHATE SYNTHASE (IPS)
NCBI Description
                  >gi 1085960 pir S52648 INO1 protein - Citrus paradisi
                  >gi 602565 emb CAA83565 (Z32632) INO1 [Citrus x paradisi]
Seq. No.
                  4477
                  3635 9.R1040
Contig ID
                  jC-gmro02910004c10a1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q4567202
BLAST score
                  294
E value
                  8.0e-35
Match length
                  127
% identity
                  93
                  (AC007168) putative myo-inositol 1-phosphate synthase
NCBI Description
                  [Arabidopsis thaliana]
Seq. No.
                  4478
Contig ID
                  3635 12.R1040
                  jC-gmle01810016g01a1
5'-most EST
                  BLASTN
Method
                  g602564
NCBI GI
BLAST score
                  87
                  4.0e-41
E value
Match length
                  261
% identity
                  83
NCBI Description C.paradisi (Macf) INO1 gene
                  4479
Seq. No.
Contig ID
                  3637 1.R1040
                  LIB3040-008-Q1-E1-H2
5'-most EST
Method
                  BLASTX
                  q3913416
NCBI GI
                  872
BLAST score
E value
                  9.0e-94
Match length
                  231
% identity
                  71
```

% identity NCBI Description

S-ADENOSYLMETHIONINE DECARBOXYLASE PROENZYME (ADOMETDC) (SAMDC) >gi_2129920_pir__S68990 adenosylmethionine

decarboxylase (EC 4.1.1.50) - Madagascar periwinkle

>gi_758695 (U12573) S-adenosyl-L-methionine decarboxylase
proenzyme [Catharanthus roseus] >gi_1094441_prf__2106177A

Met(S-adenosyl) decarboxylase [Catharanthus roseus]

Seq. No. Contig ID

3637 2.R1040

4480

5'-most EST

LIB3170-078-Q1-K1-F4



Method BLASTN
NCBI GI 9758693
BLAST score 91
E value 2.0e-43
Match length 167
% identity 89

NCBI Description Catharanthus roseus S-adenosyl-L-methionine decarboxylase

proenzyme mRNA, complete cds

Seq. No. 4481

Contig ID 3638 1.R1040

5'-most EST LIB3028-027-Q1-B2-F12

Seq. No. 4482

Contig ID 3639_1.R1040 5'-most EST leu701154652.h1

Seq. No. 4483

Contig ID 3642 1.R1040

5'-most EST LIB3028-027-Q1-B2-G10

Method BLASTN
NCBI GI g1370171
BLAST score 151
E value 2.0e-79
Match length 315
% identity 87

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB1X

Seq. No. 4484

Contig ID 3644_1.R1040

5'-most EST LIB3093-018-Q1-K1-F1

Seq. No. 4485

Contig ID 3645 1.R1040

5'-most EST LIB3051-059-Q1-K2-D9

Method BLASTX
NCBI GI g3036802
BLAST score 292
E value 5.0e-26
Match length 97
% identity 65

NCBI Description (AL022373) putative protein [Arabidopsis thaliana]

>qi 3805864 emb CAA21484 (AL031986) putative protein

[Arabidopsis thaliana]

Seq. No. 4486

Contig ID 3647_1.R1040

5'-most EST LIB3170-022-Q1-J1-E5

Method BLASTN
NCBI GI g2245073
BLAST score 42
E value 3.0e-14
Match length 66
% identity 91

NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I contig

fragment No



4487 Seq. No.

Contig ID 3648 1.R1040

5'-most EST LIB3028-008-Q1-B1-A8

BLASTX Method NCBI GI q3913733 BLAST score 679 2.0e-71 E value Match length 160 76 % identity

HYDROXYACYLGLUTATHIONE HYDROLASE CYTOPLASMIC ISOZYME NCBI Description

(GLYOXALASE II) (GLX II) >gi 1924921_emb_CAA69644_ (Y08357)

hydroxyacylglutathione hydrolase [Arabidopsis thaliana]

4488 Seq. No.

3649 1.R1040 Contig ID 5'-most EST uxk700667202.h1

Method BLASTX NCBI GI q2352492 BLAST score 1694 E value 0.0e + 00Match length 551 % identity 59

(AF005047) transport inhibitor response 1 [Arabidopsis NCBI Description

thaliana] >qi 2352494 (AF005048) transport inhibitor

response 1 [Arabidopsis thaliana]

Seq. No. 4489

Contiq ID 3649 2.R1040

LIB3139-052-P1-N1-H6 5'-most EST

Method BLASTX NCBI GI g2352492 BLAST score 143 1.0e-08 E value Match length 53 53 % identity

(AF005047) transport inhibitor response 1 [Arabidopsis NCBI Description

thaliana] >gi 2352494 (AF005048) transport inhibitor

response 1 [Arabidopsis thaliana]

4490 Seq. No.

3653 1.R1040 Contig ID

5'-most EST jC-qmro02910029d10a1

BLASTX Method g4539330 NCBI GI BLAST score 863 E value 1.0e-92 304 Match length 58 % identity

(AL035679) putative receptor-like protein kinase (fragment) NCBI Description

[Arabidopsis thaliana]

Seq. No. 4491

Contig ID 3653 2.R1040

LIB3028-027-Q1-B2-D1 5'-most EST

BLASTX Method NCBI GI q4567279 BLAST score 308



```
1.0e-27
E value
                   327
Match length
% identity
                   (ACOO6841) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   4492
                   3654 1.R1040
Contig ID
                  jC-gmro02910052a01d1
5'-most EST
                   4493
Seq. No.
                   3656 1.R1040
Contig ID
                  LIB3028-027-Q1-B2-D12
5'-most EST
Method
                  BLASTX
                   g4490706
NCBI GI
BLAST score
                   396
                   1.0e-38
E value
                   113
Match length
% identity
NCBI Description (AL035680) putative protein [Arabidopsis thaliana]
                   4494
Seq. No.
                   3657 1.R1040
Contig ID
                   LIB3107-072-Q1-K1-F2
5'-most EST
                   BLASTX
Method
                   g3522956
NCBI GI
BLAST score
                   1434
                   1.0e-159
E value
                   405
Match length
% identity
                   (AC004411) putative pectinacetylesterase precursor
NCBI Description
                   [Arabidopsis thaliana]
                   4495
Seq. No.
Contig ID
                   3657 2.R1040
                   LIB3109-010-Q1-K1-F7
5'-most EST
Method
                   BLASTX
                   g3047082
NCBI GI
BLAST score
                   453
                   4.0e-45
E value
Match length
                   132
% identity
                   67
                   (AF058914) similar to Vigna radiata pectinacetylesterase
NCBI Description
                   precursor (GB:X99348) [Arabidopsis thaliana]
                   4496
Seq. No.
                   3657 3.R1040
Contig ID
                   hyd7\overline{0}0726654.h1
5'-most EST
Method
```

Method BLASTX
NCBI GI g3522956
BLAST score 641

BLAST score 641 E value 6.0e-67 Match length 145 % identity 81

NCBI Description (AC004411) putative pectinacetylesterase precursor

[Arabidopsis thaliana]



4497 Seq. No. 3658 1.R1040 Contig ID LIB3139-010-P1-N1-B5 5'-most EST Method BLASTX q4417289 NCBI GI BLAST score 530 9.0e-54 E value 197 Match length % identity (AC007019) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 3658 2.R1040 Contig ID 5'-most EST fC-gmle700875159r1 BLASTX Method NCBI GI g4417289 BLAST score 263 8.0e-23 E value Match length 73 62 % identity NCBI Description (AC007019) unknown protein [Arabidopsis thaliana] Seq. No. 4499 3660 1.R1040 Contig ID LIB3049-041-Q1-E1-G4 5'-most EST BLASTX Method NCBI GI g2160166 BLAST score 522 5.0e-53 E value Match length 208 % identity (AC000132) No definition line found [Arabidopsis thaliana] NCBI Description Seq. No. 4500 3662 1.R1040 Contig ID LIB3170-012-Q1-K1-C5 5'-most EST Method BLASTX g464621 NCBI GI BLAST score 581

2.0e-79 E value Match length 194 80 % identity

60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi_280374_pir__S28586 NCBI Description

ribosomal protein ML16 - common ice plant

>qi 19539 emb CAA49175 (X69378) ribosomal protein YL16

[Mesembryanthemum crystallinum]

4501 Seq. No.

3662 2.R1040 Contig ID 5'-most EST fua701040224.hl

Method BLASTX NCBI GI g3860277 BLAST score 874 E value 4.0e-94 Match length 215 % identity 80

(AC005824) putative ribosomal protein L10 [Arabidopsis NCBI Description



thaliana] >gi_4314394_gb_AAD15604_ (AC006232) putative ribosomal protein L10A [Arabidopsis thaliana]

Seq. No. 4502

Contig ID 3662_4.R1040

5'-most EST LIB3094-017-Q1-K1-H8

Method BLASTX
NCBI GI g464621
BLAST score 231
E value 7.0e-38
Match length 101
% identity 80

NCBI Description 60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi_280374_pir__S28586

ribosomal protein ML16 - common ice plant

>gi 19539 emb_CAA49175_ (X69378) ribosomal protein YL16

[Mesembryanthemum crystallinum]

Seq. No. 4503

Contig ID 3662_5.R1040 5'-most EST hyd700726772.h1

Method BLASTX
NCBI GI g3860277
BLAST score 444
E value 6.0e-44
Match length 117
% identity 86

NCBI Description (AC005824) putative ribosomal protein L10 [Arabidopsis

thaliana] >gi_4314394_gb_AAD15604_ (AC006232) putative

ribosomal protein L10A [Arabidopsis thaliana]

Seq. No. 4504

3662 7.R1040 Contig ID 5'-most EST q5510388 BLASTX Method g464621 NCBI GI BLAST score 247 4.0e-21 E value Match length 56 88 % identity

NCBI Description 60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi_280374_pir__S28586

ribosomal protein ML16 - common ice plant

>gi 19539 emb CAA49175 (X69378) ribosomal protein YL16

[Mesembryanthemum crystallinum]

Seq. No. 4505

Contig ID 3662_9.R1040 5'-most EST uxk700672494.h1

Method BLASTX
NCBI GI g1709970
BLAST score 375
E value 8.0e-36
Match length 87
% identity 84

NCBI Description 60S RIBOSOMAL PROTEIN L10A

Seq. No. 4506

Contig ID 3662_12.R1040

Match length

% identity

139

41



```
LIB3053-003-Q1-N1-D8
5'-most EST
                  4507
Seq. No.
                  3664 1.R1040
Contig ID
5'-most EST
                  LIB3030-005-Q1-B1-H1
                  4508
Seq. No.
                  3671 1.R1040
Contig ID
                  LIB3028-027-Q1-B2-B11
5'-most EST
Method
                  BLASTX
                  q3461839
NCBI GI
                  437
BLAST score
                  2.0e-43
E value
                  129
Match length
% identity
                  (AC005315) putative receptor protein kinase [Arabidopsis
NCBI Description
                  thaliana]
                  4509
Seq. No.
                  3672 1.R1040
Contig ID
                  LIB3039-026-Q1-E1-C4
5'-most EST
Method
                  BLASTX
                  q2582665
NCBI GI
                  1284
BLAST score
                   1.0e-142
E value
                   282
Match length
% identity
                   88
NCBI Description (Z82983) thi [Citrus sinensis]
Seq. No.
Contig ID
                   3673 1.R1040
                   LIB3028-027-Q1-B2-B2
5'-most EST
                   4511
Seq. No.
                   3674 1.R1040
Contig ID
                   pcp700994538.h2
5'-most EST
                   4512
Seq. No.
                   3676 1.R1040
Contig ID
                   hrw701060639.h1
5'-most EST
Method
                   BLASTX
                   q3063467
NCBI GI
BLAST score
                   467
                   2.0e-46
E value
                   126
Match length
                   73
% identity
                  (AC003981) F22013.29 [Arabidopsis thaliana]
NCBI Description
                   4513
Seq. No.
                   3682 1.R1040
Contig ID
5'-most EST
                   fde700873509.hl
Method
                   BLASTX
                   q4263818
NCBI GI
BLAST score
                   189
                   4.0e-15
E value
```



NCBI Description (AC006067) unknown protein [Arabidopsis thaliana]

Seq. No. 4514

Contig ID 3694 1.R1040

5'-most EST LIB3039-022-Q1-E1-D6

Seq. No. 4515

Contig ID 3694 2.R1040

5'-most EST LIB3106-092-Q1-K1-G8

Seq. No. 4516

Contig ID 3694_4.R1040

5'-most EST LIB3049-017-Q1-E1-H1

Seq. No. 4517

Contig ID 3695_1.R1040

5'-most EST LIB3028-027-Q1-B2-A10

Seq. No. 4518

Contig ID 3697_1.R1040

5'-most EST LIB3028-027-Q1-B1-A11

Seq. No. 4519

Contig ID 3700_1.R1040 5'-most EST hyd700728205.h1

Seq. No. 4520

Contig ID 3700_2.R1040

5'-most EST LIB3050-024-Q1-K1-A2

Seq. No. 4521

Contig ID 3701_1.R1040 5'-most EST epx701109853.h1

Method BLASTX
NCBI GI g2136139
BLAST score 168
E value 9.0e-12
Match length 144
% identity 8

NCBI Description sds22 protein homolog - human >gi_1085028_emb_CAA90626_

(Z50749) yeast sds22 homolog [Homo sapiens]

>gi_1585165_prf__2124310A sds22 gene [Homo sapiens]

>gi_4506013_ref_NP_002703.1_pPPPP1R7_ protein phosphatase 1,

regulatory subunit

Seq. No. 4522

Contig ID 3712 1.R1040

5'-most EST LIB3049-033-Q1-E1-G7

Method BLASTN
NCBI GI g2565428
BLAST score 48
E value 1.0e-17
Match length 112
% identity 86

NCBI Description Onobrychis viciifolia glycine-rich protein mRNA, complete

cds



4523 Seq. No.

3712 3.R1040 Contig ID

LIB3073-026-Q1-K1-D8 5'-most EST

Method BLASTN NCBI GI g2565428 BLAST score 46 9.0e-17 E value Match length 106 % identity

Onobrychis viciifolia glycine-rich protein mRNA, complete NCBI Description

Seq. No. 4524

3715 1.R1040 Contig ID g5753666 5'-most EST BLASTN Method NCBI GI g408793 BLAST score 472 E value 0.0e + 00613 Match length 96 % identity

Glycine soja chloroplast 3-omega faty acid desaturase NCBI Description

(Fad3) mRNA, complete cds

4525 Seq. No.

3715 2.R1040 Contig ID 5'-most EST leu701149421.h1

BLASTN Method g408793 NCBI GI BLAST score 375 0.0e + 00E value 379 Match length % identity 100

Glycine soja chloroplast 3-omega faty acid desaturase NCBI Description

(Fad3) mRNA, complete cds

4526 Seq. No.

Contig ID 3716 1.R1040

5'-most EST jC-gmf102220061a11a1

Method BLASTX NCBI GI q1709449 1754 BLAST score E value 0.0e + 00396 Match length % identity 85

PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT NCBI Description

PRECURSOR (PDHE1-A) >gi_1263302 (U51918) pyruvate

dehydrogenase E1 alpha subunit [Pisum sativum]

4527 Seq. No.

3716 2.R1040 Contig ID

5'-most EST LIB3170-041-Q1-K1-C9

Method BLASTX g1709449 NCBI GI 1118 BLAST score 1.0e-175 E value Match length 370



% identity PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT NCBI Description PRECURSOR (PDHE1-A) >gi_1263302 (U51918) pyruvate dehydrogenase E1 alpha subunit [Pisum sativum] Seq. No. 4528 Contig ID 3716 3.R1040 5'-most EST fC-gmse700756337a2 BLASTN Method q3851004 106 2.0e-52

NCBI GI BLAST score E value 194 Match length 89 % identity

Zea mays pyruvate dehydrogenase E1 alpha subunit RNA, NCBI Description nuclear gene encoding mitochondrial protein, complete cds

Seq. No. Contig ID

3716 6.R1040

5'-most EST uC-gmropic041f09b1

4530 Seq. No.

3721 1.R1040 Contig ID

jC-gmle01810094c05a1 5'-most EST

BLASTX Method NCBI GI g4455246 BLAST score 691 1.0e-72 E value 238 Match length % identity

(AL035523) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 4531

3721 2.R1040 Contig ID

LIB3138-031-Q1-N1-D2 5'-most EST

Method BLASTX NCBI GI g4455246 BLAST score 334 3.0e-31 E value Match length 95 % identity 69

(AL035523) putative protein [Arabidopsis thaliana] NCBI Description

4532 Seq. No.

3721 3.R1040 Contig ID fua701043571.h1 5'-most EST

Method BLASTX g4455246 NCBI GI 191 BLAST score E value 9.0e-15 Match length 62 % identity 75

(AL035523) putative protein [Arabidopsis thaliana] NCBI Description

4533 Seq. No.

Contig ID 3723 1.R1040 jex700906481.h1 5'-most EST



```
BLASTX
Method
NCBI GI
                   g3176098
                   722
BLAST score
                   2.0e-76
E value
                   191
Match length
                   47
% identity
NCBI Description (Y15036) annexin [Medicago truncatula]
Seq. No.
                   4534
                   3723 2.R1040
Contig ID
                   ek17\overline{0}0968268.h1
5'-most EST
                   BLASTX
Method
                   g3176098
NCBI GI
                   244
BLAST score
                   5.0e-21
E value
Match length
                   64
                   77
% identity
NCBI Description (Y15036) annexin [Medicago truncatula]
                   4535
Seq. No.
                   3724 1.R1040
Contig ID
                   kl1701210733.h1
5'-most EST
                   BLASTX
Method
                   q1666096
NCBI GI
BLAST score
                   1012
                   1.0e-110
E value
                   297
Match length
                   68
% identity
NCBI Description (Y09113) dioxygenase [Marah macrocarpus]
                    4536
Seq. No.
                   3724 2.R1040
Contig ID
                   sat7\overline{0}1005739.h1
5'-most EST
Method
                   BLASTX
                    g1666096
NCBI GI
BLAST score
                   171
                    2.0e-12
E value
Match length
                    61
% identity
                    59
NCBI Description (Y09113) dioxygenase [Marah macrocarpus]
Seq. No.
Contig ID
                    3725 1.R1040
                    uC-gmflminsoy026e06b1
5'-most EST
                    BLASTX
Method
NCBI GI
                    q2194118
BLAST score
                    222
E value
                    6.0e-18
Match length
                    74
% identity
                    46
                   (AC002062) F20P5.4 gene product [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    4538
                    3725 2.R1040
Contig ID
                    k117\overline{0}1211135.h1
5'-most EST
```

884

BLASTX

q2194118

Method NCBI GI



BLAST score 221 E value 1.0e-17 Match length 95 % identity 40

NCBI Description (AC002062) F20P5.4 gene product [Arabidopsis thaliana]

Seq. No. 4539

Contig ID 3726 1.R1040

5'-most EST jC-gmf102220057d05a1

Method BLASTX
NCBI GI g3881189
BLAST score 573
E value 5.0e-59
Match length 155
% identity 68

NCBI Description (Z99281) similar to ADP-ribosylation factor; cDNA EST

EMBL:C08179 comes from this gene; cDNA EST EMBL:C08337 comes from this gene; cDNA EST EMBL:C09829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4

Seq. No. 4540

Contig ID 3726_2.R1040 5'-most EST pmv700889323.h1

Method BLASTX
NCBI GI g3881189
BLAST score 345
E value 1.0e-32
Match length 77
% identity 78

NCBI Description (Z99281) similar to ADP-ribosylation factor; cDNA EST

EMBL:C08179 comes from this gene; cDNA EST EMBL:C08337 comes from this gene; cDNA EST EMBL:C09829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4

Seq. No. 4541

Contig ID 3727_1.R1040

5'-most EST LIB3039-017-Q1-E1-A3

Seq. No. 4542

Contig ID 3729_1.R1040 5'-most EST wvk700686507.h1

Method BLASTN
NCBI GI g2618601
BLAST score 35
E value 2.0e-09
Match length 87
% identity 85

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MHJ24, complete sequence [Arabidopsis thaliana]

Seq. No. 4543

Contig ID 3730_1.R1040

5'-most EST LIB3049-005-Q1-E1-C1

Method BLASTX
NCBI GI g4455364
BLAST score 817
E value 2.0e-87



```
Match length
% identity
                   (AL035524) senescence-associated protein-like [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                   4544
                   3730 2.R1040
Contig ID
                  epx701109821.hl
5'-most EST
                  BLASTX
Method
                   g3551954
NCBI GI
BLAST score
                   477
                   2.0e-47
E value
                   218
Match length
                   47
% identity
                   (AF082030) senescence-associated protein 5 [Hemerocallis
NCBI Description
                   hybrid cultivar]
                   4545
Seq. No.
                   3730 4.R1040
Contig ID
                   uC-gmflminsoy063e02b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4455364
                   354
BLAST score
                   1.0e-33
E value
                   97
Match length
                   68
% identity
                   (AL035524) senescence-associated protein-like [Arabidopsis
NCBI Description
                   thaliana]
                   4546
Seq. No.
                   3732 1.R1040
Contig ID
                   uC-qmropic050g09b1
5'-most EST
                   4547
Seq. No.
                   3733 1.R1040
Contig ID
                   LIB3051-016-Q1-E1-C12
5'-most EST
Method
                   BLASTN
                   g3821780
NCBI GI
BLAST score
                   33
                   6.0e-09
E value
Match length
                   37
                   97
% identity
NCBI Description Xenopus laevis cDNA clone 27A6-1
                   4548
Seq. No.
                   3742 1.R1040
Contig ID
                   LIB3050-002-Q1-E1-E7
5'-most EST
Method
                   BLASTX
                   g2827528
NCBI GI
                   383
BLAST score
                   6.0e-37
E value
                   98
Match length
% identity
                   76
                   (AL021633) predicted protein [Arabidopsis thaliana]
NCBI Description
```

4549

3743_1.R1040

Seq. No.

Contig ID



```
LIB3049-038-Q1-E1-C6
5'-most EST
                  BLASTN
Method
                  q3413169
NCBI GI
BLAST score
                   345
                  0.0e + 00
E value
                  577
Match length
                   90
% identity
NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S6, partial
                   4550
Seq. No.
                   3743 2.R1040
Contig ID
                   LIB3039-019-Q1-E1-F7
5'-most EST
                   BLASTN
Method
                   g3413169
NCBI GI
BLAST score
                   231
                   1.0e-127
E value
                   407
Match length
                   89
% identity
NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S6, partial
                   4551
Seq. No.
                   3743 3.R1040
Contia ID
                   LIB3\overline{0}28-026-Q1-B1-F10
5'-most EST
Method
                   BLASTN
                   g3413169
NCBI GI
BLAST score
                   137
                   6.0e-71
E value
Match length
                   229
                   90
% identity
                   Cicer arietinum mRNA for 40S ribosomal protein S6, partial
NCBI Description
                   4552
Seq. No.
                   3743 5.R1040
Contig ID
                   LIB3051-020-Q1-E1-H7
5'-most EST
                   BLASTN
Method
                   g3413169
NCBI GI
BLAST score
                   261
                   1.0e-145
E value
                   457
Match length
                   89
% identity
NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S6, partial
                   4553
Seq. No.
                   3743 6.R1040
Contig ID
                   LIB3040-029-Q1-E1-F8
5'-most EST
                   4554
Seq. No.
                   3746 1.R1040
Contig ID
                   LIB3040-041-Q1-E1-D4
5'-most EST
                   BLASTN
Method
                   g2879810
NCBI GI
BLAST score
                   155
                   2.0e-81
E value
                   319
Match length
                   87
 % identity
```

NCBI Description Lupinus luteus mRNA for ribosomal protein L30



```
4555
Seq. No.
Contig ID
                  3746 2.R1040
5'-most EST
                  LIB3073-012-Q1-K1-C3
                  BLASTN
Method
NCBI GI
                  g2879810
BLAST score
                  169
                  5.0e-90
E value
                  341
Match length
                  87
% identity
                  Lupinus luteus mRNA for ribosomal protein L30
NCBI Description
Seq. No.
                   4556
                  3747 1.R1040
Contig ID
5'-most EST
                  txt700735453.h1
                  BLASTX
Method
NCBI GI
                  q3402683
BLAST score
                  563
E value
                   1.0e-57
                  266
Match length
                   44
% identity
                  (AC004697) patatin-like protein [Arabidopsis thaliana]
NCBI Description
                   4557
Seq. No.
                   3747 2.R1040
Contig ID
5'-most EST
                   jC-gmst02400077e07d1
Seq. No.
                   4558
                   3747 4.R1040
Contig ID
5'-most EST
                  LIB3039-037-Q1-E1-B12
                   4559
Seq. No.
Contig ID
                   3748 1.R1040
5'-most EST
                   awf700837610.hl
Seq. No.
                   4560
                   3748 2.R1040
Contig ID
5'-most EST
                   uC-gmropic008a11b1
                   4561
Seq. No.
                   3748 3.R1040
Contig ID
5'-most EST
                   kmv700743318.h1
                   4562
Seq. No.
                   3749 1.R1040
Contig ID
                   leu701154836.h1
5'-most EST
                   BLASTX
Method
                   g2160182
NCBI GI
BLAST score
                   264
```

8.0e-23 E value 146 Match length % identity 40

NCBI Description

(AC000132) ESTs gb_ATTS1236, gb_T43334, gb_N97019, gb_AA395203

come from this gene. [Arabidopsis thaliana]

Seq. No. Contig ID 4563

3750 1.R1040

5'-most EST

LIB3109-013-Q1-K1-E2



Method BLASTX
NCBI GI g731507
BLAST score 179
E value 2.0e-12
Match length 254
% identity 26

NCBI Description HYPOTHETICAL 48.8 KD PROTEIN IN SSU81-SCS2 INTERGENIC

REGION >gi 1077700_pir__S50622 hypothetical protein YER119c

- yeast (Saccharomyces cerevisiae) >gi_603358 (U18916)

Yer119cp [Saccharomyces cerevisiae]

Seq. No. 4564

Contig ID 3751 1.R1040

5'-most EST LIB3028-026-Q1-B1-C6

Method BLASTX
NCBI GI g4262174
BLAST score 1482
E value 1.0e-165
Match length 398
% identity 70

NCBI Description (AC005508) 9058 [Arabidopsis thaliana]

Seq. No. 4565

3753 1.R1040 Contig ID g4397335 5'-most EST BLASTX Method NCBI GI g3319355 BLAST score 1481 1.0e-165 E value 314 Match length 89 % identity

NCBI Description (AF077407) similar to chaperonin containing TCP-1 complex

gamma chain [Arabidopsis thaliana]

Seq. No. 4566

Contig ID 3753 2.R1040 5'-most EST kl1701207044.h1

Method BLASTX
NCBI GI g3319355
BLAST score 538
E value 7.0e-64
Match length 281
% identity 72

NCBI Description (AF077407) similar to chaperonin containing TCP-1 complex

gamma chain [Arabidopsis thaliana]

Seq. No. 4567

Contig ID 3753 3.R1040

5'-most EST LIB3049-021-Q1-E1-H7

Method BLASTX
NCBI GI g3319355
BLAST score 789
E value 2.0e-84
Match length 172
% identity 87

NCBI Description (AF077407) similar to chaperonin containing TCP-1 complex

gamma chain [Arabidopsis thaliana]



Seq. No. 4568

Contig ID 3753_5.R1040 5'-most EST zzp700834785.h1

Method BLASTX
NCBI GI g3319355
BLAST score 250
E value 2.0e-21
Match length 55
% identity 85

NCBI Description (AF077407) similar to chaperonin containing TCP-1 complex

gamma chain [Arabidopsis thaliana]

Seq. No. 4569

Contig ID 3753_6.R1040 5'-most EST uxk700669084.h1

Seq. No. 4570

Contig ID 3754 1.R1040

5'-most EST LIB3051-037-Q1-K1-H10

Method BLASTX
NCBI GI g2623158
BLAST score 165
E value 3.0e-11
Match length 110
% identity 31

NCBI Description (AF030177) N-acetylglucosaminyl transferase component Gpi1

[Homo sapiens] >gi 2911142_dbj_BAA24948_ (AB003723) GPI1

[Homo sapiens]

Seq. No. 4571

Contig ID 3755_1.R1040

5'-most EST LIB3028-026-Q1-B1-D10

Method BLASTX
NCBI GI g1168314
BLAST score 357
E value 2.0e-33
Match length 235
% identity 18

NCBI Description REGULATOR OF ACETYL-COA SYNTHETASE ACTIVITY

>gi_1084627_pir__S57116 probable carrier protein ACR1 yeast (Saccharomyces cerevisiae) >gi_1015794_emb_CAA89624_

(Z49595) ORF YJR095w [Saccharomyces cerevisiae]

Seq. No. 4572

Contig ID 3756 1.R1040

5'-most EST LIB3028-026-Q1-B1-D12

Seq. No. 4573

Contig ID 3758 1.R1040

5'-most EST uC-gmflminsoy119b08b1

Method BLASTX
NCBI GI 94220527
BLAST score 943
E value 1.0e-102
Match length 342
% identity 56

Match length

143



```
NCBI Description (AL035356) putative protein [Arabidopsis thaliana]
                   4574
Seq. No.
                  3760 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220073g01a1
Method
                  BLASTX
                  g2208908
NCBI GI
                   614
BLAST score
                   6.0e-64
E value
                   132
Match length
% identity
                   (AB004809) phosphate transporter [Catharanthus roseus]
NCBI Description
Seq. No.
                   3761 1.R1040
Contig ID
5'-most EST
                   LIB3107-002-Q1-K1-D1
Method
                   BLASTX
NCBI GI
                   q3024541
BLAST score
                   163
                   8.0e-11
E value
                   293
Match length
                   6
% identity
NCBI Description RESTIN >gi_2338714 (AF014012) restin [Gallus gallus]
                   4576
Seq. No.
                   3765 1.R1040
Contig ID
5'-most EST
                   zsq701118781.hl
                   BLASTX
Method
                   q4091008
NCBI GI
BLAST score
                   310
                   5.0e-28
E value
                   101
Match length
% identity
                   (AF040700) methionyl-tRNA synthetase [Oryza sativa]
NCBI Description
                   4577
Seq. No.
Contig ID
                   3766 1.R1040
                   LIB3109-035-Q1-K1-B4
5'-most EST
                   BLASTX
Method
                   g2129927
NCBI GI
BLAST score
                   2238
                   0.0e + 00
E value
                   487
Match length
                   87
% identity
                   zeta-carotene desaturase precursor - pepper
NCBI Description
                   >gi_1176437_bbs_171885 zeta-carotene desaturase,
                   CapZDS=phytoene desaturase homolog [Capsicum annuum, early
                   ripening fruit, Peptide, 588 aa]
                   4578
Seq. No.
                   3771 1.R1040
Contig ID
                   pxt700945660.h1
5'-most EST
Method
                   BLASTX
                   q2459429
NCBI GI
                   599
BLAST score
E value
                   4.0e-62
```



% identity NCBI Description (AC002332) unknown protein [Arabidopsis thaliana]

4579 Seq. No.

3771 2.R1040 Contig ID

5'-most EST LIB3092-032-Q1-K1-F8

Method BLASTX g2459429 NCBI GI 580 BLAST score 7.0e-60 E value 147 Match length 74 % identity

NCBI Description (AC002332) unknown protein [Arabidopsis thaliana]

Seq. No.

4580 3772 1.R1040 Contig ID

5'-most EST LIB3093-037-Q1-K1-E4

BLASTX Method NCBI GI g3790587 996 BLAST score 1.0e-108 E value 347

Match length 61 % identity

(AF079182) RING-H2 finger protein RHF2a [Arabidopsis NCBI Description

thaliana]

4581 Seq. No.

3772 2.R1040 Contig ID rca700997363.h1 5'-most EST

BLASTX Method a3790587 NCBI GI 360 BLAST score E value 5.0e - 3474 Match length % identity 88

(AF079182) RING-H2 finger protein RHF2a [Arabidopsis NCBI Description

thaliana]

Seq. No. 4582

3777 1.R1040 Contig ID zhf700953648.h1 5'-most EST

BLASTX Method q3877358 NCBI GI 188 BLAST score 6.0e-14E value 85 Match length 28 % identity

(Z66520) similar to RBB3 like protein; cDNA EST EMBL:C08891 NCBI Description

comes from this gene; cDNA EST EMBL:C09371 comes from this

gene; cDNA EST yk468f10.5 comes from this gene

[Caenorhabditis elegans]

4583 Seq. No.

3778 1.R1040 Contig ID smc700745972.h1 5'-most EST

Method BLASTX g2443751 NCBI GI



```
BLAST score
                  1368
                  1.0e-162
E value
                  351
Match length
                  84
% identity
                  (AF020303) fumarase [Arabidopsis thaliana] >gi 2529676
NCBI Description
                   (AC002535) putative fumarase [Arabidopsis thaliana]
Seq. No.
                   3778 2.R1040
Contig ID
                  asn701138372.h1
5'-most EST
                  BLASTX
Method
                   g2443751
NCBI GI
                   292
BLAST score
                   3.0e-26
E value
                   60
Match length
                   92
% identity
                   (AF020303) fumarase [Arabidopsis thaliana] >gi_2529676
NCBI Description
                   (AC002535) putative fumarase [Arabidopsis thaliana]
Seq. No.
                   3778 3.R1040
Contig ID
                   leu7\overline{0}1157217.h1
5'-most EST
                   BLASTX
Method
                   g2443751
NCBI GI
BLAST score
                   146
                   2.0e-09
E value
                   61
Match length
% identity
                   (AF020303) fumarase [Arabidopsis thaliana] >gi 2529676
NCBI Description
                   (AC002535) putative fumarase [Arabidopsis thal\overline{i}ana]
Seq. No.
Contig ID
                   3779 1.R1040
                   epx701108935.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1076534
BLAST score
                   1833
                   0.0e + 00
E value
                   433
Match length
                   81
% identity
                   monodehydroascorbate reductase (NADH) (EC 1.6.5.4) - garden
NCBI Description
                   pea >gi_497120 (U06461) monodehydroascorbate reductase
                   [Pisum sativum]
                   4587
Seq. No.
                   3784 1.R1040
Contig ID
                   LIB3093-032-Q1-K1-C6
5'-most EST
Method
                   BLASTX
                   g2245378
NCBI GI
                   819
BLAST score
                   3.0e-87
E value
Match length
                   398
% identity
                   56
                   (U83245) auxin response factor 1 [Arabidopsis thaliana]
NCBI Description
```

893

4588

3786_1.R1040

Seq. No.

Contig ID



5'-most EST LIB3028-026-Q1-B1-A10

Method BLASTX
NCBI GI g586021
BLAST score 218
E value 2.0e-17
Match length 96
% identity 45

NCBI Description PROBABLE PEPTIDYL-TRNA HYDROLASE (PTH) (STAGE V SPORULATION

PROTEIN C) >gi_2127242_pir__S66083 stage V sporulation protein - Bacillus subtilis >gi_467442_dbj_BAA05288_(D26185) stage V sporulation [Bacillus subtilis]

>gi 2632320 emb CAB11829 (Z99104) thermosensitive mutant blocks spore coat formation (stage V sporulation) [Bacillus

subtilis]

Seq. No. 4589

Contig ID 3787_1.R1040

5'-most EST LIB3028-026-Q1-B1-A11

Seq. No. 4590

Contig ID 3789_1.R1040 5'-most EST jex700908228.h1

Method BLASTX
NCBI GI g3256770
BLAST score 306
E value 3.0e-27
Match length 314
% identity 31

NCBI Description (AP000002) 318aa long hypothetical UDP-glucose 4-epimerase

[Pyrococcus horikoshii]

Seq. No. 4591

Contig ID 3789_3.R1040 5'-most EST trc700567883.h1

Seq. No. 4592

Contig ID 3789 4.R1040 5'-most EST ncj700987768.h1

Seq. No. 4593

Contig ID 3789_6.R1040

5'-most EST uC-gmronoir031f04b1

Seq. No. 4594

Contig ID 3794 1.R1040

5'-most EST LIB3028-025-Q1-B1-H11

Method BLASTX
NCBI GI g3461817
BLAST score 287
E value 7.0e-26
Match length 66
% identity 80

NCBI Description (AC004138) unknown protein [Arabidopsis thaliana]

Seq. No. 4595

Contig ID 3795 1.R1040

5'-most EST uC-gmflminsoy057d10b1



```
ren.
 Method
                    BLASTN
 NCBI GI
                    g556421
 BLAST score
                    45
                    1.0e-15
 E value
                    247
 Match length
                    89
 % identity
                    Stylosanthes humilis cinnamyl alcohol dehydrogenase (CAD1)
 NCBI Description
                    mRNA, complete cds
                    4596
 Seq. No.
                    3795 2.R1040
 Contig ID
                    LIB3028-025-Q1-B1-F5
 5'-most EST
                    BLASTX
 Method
 NCBI GI
                    g4056499
 BLAST score
                    149
                    2.0e-09
 E value
 Match length
                    44
                    66
 % identity
                   (AC005896) unknown protein [Arabidopsis thaliana]
 NCBI Description
                    4597
 Seq. No.
 Contig ID
                    3798 1.R1040
                    ncj7\overline{0}0983420.h1
 5'-most EST
                    4598
 Seq. No.
                    3798 2.R1040
 Contig ID
 5'-most EST
                    ncj700978215.h1
                    4599
 Seq. No.
 Contig ID
                    3803 1.R1040
 5'-most EST
                    hyd700726639.h1
                    BLASTX
 Method
 NCBI GI
                    q4539383
 BLAST score
                    535
 E value
                    1.0e-54
 Match length
                    189
 % identity
                    54
                    (AL035526) putative protein (fragment) [Arabidopsis
 NCBI Description
                    thaliana]
 Seq. No.
                    4600
                    3806 1.R1040
 Contig ID
 5'-most EST
                    leu701153143.h1
 Method
                    BLASTX
                    q4115377
 NCBI GI
 BLAST score
                    664
 E value
                    2.0e-72
                    292
 Match length
 % identity
                    45
 NCBI Description
                    (AC005967) unknown protein [Arabidopsis thaliana]
 Seq. No.
                    4601
                    3806 2.R1040
 Contig ID
```

5'-most EST jC-gmst02400046c12d1

4602 Seq. No.

Contig ID 3808 1.R1040



```
zzp700831734.hl
5'-most EST
                  BLASTX
Method
                  g3341697
NCBI GI
                   290
BLAST score
                   3.0e-53
E value
                  143
Match length
% identity
                   (AC003672) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   3809 1.R1040
Contig ID
                   eep700867020.hl
5'-most EST
                   BLASTX
Method
                   g3386597
NCBI GI
                   222
BLAST score
                   7.0e-18
E value
                   77
Match length
                   61
% identity
                   (AC004665) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi_3702347 (AC005397) putative permease [Arabidopsis
                   thaliana]
                   4604
Seq. No.
                   3809 2.R1040
Contig ID
                   xpa700793883.hl
5'-most EST
                   4605
Seq. No.
                   3810 1.R1040
Contig ID
                   pxt700946473.h1
5'-most EST
                   BLASTX
Method
                   q3461835
NCBI GI
                   1325
BLAST score
                   1.0e-146
E value
                   408
Match length
% identity
                   (AC005315) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   >gi 3927840 (AC005727) putative protein kinase [Arabidopsis
                   thaliana]
                    4606
 Seq. No.
                    3810 3.R1040
 Contig ID
                   taw700657682.h1
 5'-most EST
                    4607
 Seq. No.
                    3812 1.R1040
 Contig ID
                    sat701013536.hl
 5'-most EST
                    BLASTX
 Method
                    g3123515
 NCBI GI
                    713
 BLAST score
                    3.0e-75
 E value
                    146
 Match length
                    92
 % identity
```

NCBI Description (Y08761) Mago Nashi-like protein [Euphorbia lagascae]

Sea. No. 4608

Seq. No. 4608 Contig ID 3812_2.R1040 5'-most EST ncj700977319.h1



```
BLASTN
Method
NCBI GI
                  g3006211
BLAST score
                  73
                  6.0e-33
E value
Match length
                  133
% identity
                  89
                  Drosophila melanogaster (P1 DS02368 (D205)) DNA sequence,
NCBI Description
                  complete sequence [Drosophila melanogaster]
                   4609
Seq. No.
Contig ID
                   3813 1.R1040
                  rca700996410.h1
5'-most EST
                   4610
Seq. No.
                   3813 2.R1040
Contig ID
5'-most EST
                   LIB3040-041-Q1-E1-D1
Method
                   BLASTX
NCBI GI
                   q3115852
BLAST score
                   460
                   1.0e-45
E value
                   107
Match length
% identity
                   80
                  (AL022023) putative protein [Arabidopsis thaliana]
NCBI Description
                   4611
Seq. No.
                   3813 3.R1040
Contig ID
5'-most EST
                   q4277052
                   BLASTX
Method
                   q3115852
NCBI GI
BLAST score
                   356
E value
                   9.0e-34
Match length
                   71
% identity
                   92
                  (AL022023) putative protein [Arabidopsis thaliana]
NCBI Description
                   4612
Seq. No.
                   3814 1.R1040
Contig ID
                   LIB3139-055-P1-N1-G5
5'-most EST
                   BLASTX
Method
                   q3738283
NCBI GI
BLAST score
                   421
                   3.0e-41
E value
                   83
Match length
                   89
% identity
                  (AC005309) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4613
Seq. No.
                   3815_1.R1040
Contig ID
                   jC-gmf102220085d03a1
5'-most EST
Method
                   BLASTX
NCBI GI
                   a886132
                   408
BLAST score
```

897

(U28149) putative surface protein [Medicago sativa]

1.0e-39

235

E value

Match length % identity

NCBI Description



Seq. No. 4614

Contig ID 3815_2.R1040

5'-most EST LIB3028-025-Q1-B1-E5

Method BLASTX
NCBI GI g886132
BLAST score 252
E value 1.0e-21
Match length 97
% identity 52

NCBI Description (U28149) putative surface protein [Medicago sativa]

Seq. No. 4615

Contig ID 3816_1.R1040 5'-most EST bth700848583.h1

Method BLASTX
NCBI GI g2623297
BLAST score 260
E value 7.0e-22
Match length 101
% identity 52

NCBI Description (AC002409) unknown protein [Arabidopsis thaliana]

>gi 3790583 (AF079180) RING-H2 finger protein RHC1a

[Arabidopsis thaliana]

Seq. No. 4616

Contig ID 3816_2.R1040 5'-most EST epx701108636.h1

Method BLASTX
NCBI GI g3128178
BLAST score 201
E value 2.0e-15
Match length 58
% identity 57

NCBI Description (AC004521) hypothetical protein [Arabidopsis thaliana]

Seq. No. 4617

Contig ID 3816_3.R1040

5'-most EST LIB3028-025-Q1-B1-E6

Seq. No. 4618

Contig ID 3819 1.R1040

5'-most EST LIB3049-052-Q1-E1-D5

Method BLASTX
NCBI GI g3257202
BLAST score 209
E value 3.0e-16
Match length 121
% identity 40

NCBI Description (AP000003) 172aa long hypothetical protein [Pyrococcus

horikoshii]

Seq. No. 4619

Contig ID 3823 1.R1040

5'-most EST LIB3028-025-Q1-B1-F2

Method BLASTX
NCBI GI g3080438
BLAST score 401

8.0e-39 E value Match length 140 % identity (AL022605) putative protein [Arabidopsis thaliana] NCBI Description Seq. No. 4620 3824 1.R1040 Contig ID k11701205335.h1 5'-most EST Method BLASTN NCBI GI g287561 BLAST score 181 3.0e-97 E value 325 Match length 89 % identity NCBI Description Vigna radiata auxin-regulated mRNA Seq. No. 4621 Contig ID 3826 1.R1040 5'-most EST uC-gmropic036b10b1 BLASTX Method NCBI GI g3395432 BLAST score 1841 0.0e+00E value 450 Match length % identity (AC004683) unknown protein [Arabidopsis thaliana] NCBI Description 4622 Seq. No. Contig ID 3826 2.R1040 2DC-01-Q1-E1-E1 5'-most EST Method BLASTX NCBI GI q3915826 1124 BLAST score 1.0e-123 E value 294 Match length 72 % identity 60S RIBOSOMAL PROTEIN L5 NCBI Description

Seq. No. 4623

Contig ID 3826_3.R1040

5'-most EST fC-gmse700654972d1

Method BLASTX
NCBI GI g3395432
BLAST score 278
E value 1.0e-24
Match length 65
% identity 77

NCBI Description (AC004683) unknown protein [Arabidopsis thaliana]

Seq. No. 4624

Contig ID 3826_4.R1040

5'-most EST jC-gmst02400017g01d1

Method BLASTX
NCBI GI g3914771
BLAST score 422
E value 2.0e-41
Match length 101



```
% identity
                  60S RIBOSOMAL PROTEIN L5 >gi 1881380_dbj_BAA19415_
NCBI Description
                  (AB001583) ribosomal protein L5 [Solanum melongena]
                  4625
Seq. No.
Contig ID
                  3826 5.R1040
                  q5753460
5'-most EST
                  BLASTX
Method
                  q3915826
NCBI GI
                  575
BLAST score
E value
                  3.0e-59
                  140
Match length
% identity
                  60S RIBOSOMAL PROTEIN L5
NCBI Description
                  4626
Seq. No.
                  3826 6.R1040
Contig ID
5'-most EST
                  LIB3094-047-Q1-K1-D3
                  BLASTX
Method
                  q3915826
NCBI GI
BLAST score
                  534
E value
                  1.0e-54
                 · 121
Match length
                  82
% identity
NCBI Description 60S RIBOSOMAL PROTEIN L5
Seq. No.
                   4627
                   3826 8.R1040
Contig ID
                   uC-gmrominsoy121e11b1
5'-most EST
Method
                   BLASTN
                   g1183936
NCBI GI
                   64
BLAST score
E value
                   2.0e-27
                   119
Match length
                   89
% identity
NCBI Description P.sativum 5S rRNA gene
Seq. No.
                   4628
                   3826 12.R1040
Contig ID
                   awf700841854.hl
5'-most EST
Method
                   BLASTX
                   q3914771
NCBI GI
BLAST score
                   185
```

1.0e-13 E value Match length 42 % identity 86

60S RIBOSOMAL PROTEIN L5 >gi_1881380_dbj_BAA19415 NCBI Description (AB001583) ribosomal protein L5 [Solanum melongena]

4629 Seq. No.

3826 13.R1040 Contig ID $jex7\overline{0}0908485.h1$ 5'-most EST

Method BLASTX g3395432 NCBI GI 332 BLAST score 7.0e-31 E value 106 Match length



% identity NCBI Description (AC004683) unknown protein [Arabidopsis thaliana] 4630 Seq. No. 3826 19.R1040 Contig ID 5'-most EST LIB3029-008-Q1-B1-G1 Method BLASTX NCBI GI g3914771 253 BLAST score 1.0e-21 E value 67 Match length 73 % identity 60S RIBOSOMAL PROTEIN L5 >gi 1881380_dbj_BAA19415_ NCBI Description (AB001583) ribosomal protein L5 [Solanum melongena] 4631 Seq. No. Contig ID 3826 20.R1040 5'-most EST LIB3106-022-Q1-K1-E7 Seq. No. 4632 3829 1.R1040 Contig ID LIB3039-020-Q1-E1-C4 5'-most EST Method BLASTN q2565428 NCBI GI BLAST score 41 1.0e-13 E value 65 Match length 91 % identity Onobrychis viciifolia glycine-rich protein mRNA, complete NCBI Description cds 4633 Seq. No. Contig ID 3836 1.R1040 jC-gmst02400064a12d1 5'-most EST 4634 Seq. No. Contig ID 3838 1.R1040 5'-most EST LIB3170-054-Q1-J1-D11 Method BLASTX NCBI GI q2462835 BLAST score 180 E value 5.0e-13 Match length 104 % identity 32 NCBI Description (AF000657) hypothetical protein [Arabidopsis thaliana] Seq. No. 4635 3838 2.R1040 Contig ID 5'-most EST LIB3028-025-Q1-B1-D3

Seq. No. Contig ID

3840_1.R1040

4636

E value 5.0e-69



Match length 361 % identity NCBI Description Cicer arietinum mRNA for plantacyanin

Seq. No. 4637

Contig ID 3844 1.R1040

LIB3170-079-Q1-K1-C7 5'-most EST

BLASTX Method g4220461 NCBI GI BLAST score 147 E value 3.0e-09 73 Match length 45 % identity

NCBI Description (AC006216) ESTs gb T75642 and gb AA650997 come from this

gene. [Arabidopsis thaliana]

Seq. No. 4638

Contig ID 3845 1.R1040

5'-most EST LIB3170-023-Q1-K1-F9

4639 Seq. No.

Contig ID 3845 2.R1040

5'-most EST LIB3039-032-Q1-E1-F2

4640 Seq. No.

3846 1.R1040 Contig ID 5'-most EST 6HA - 02 - Q1 - E1 - A7

4641 Seq. No.

Contig ID 3848 1.R1040

5'-most EST LIB3050-024-Q1-K1-C6

Method BLASTX g2500341 NCBI GI BLAST score 165 5.0e-11 E value Match length 38 % identity 74

50S RIBOSOMAL PROTEIN L36 >gi_1652406_dbj_BAA17328 NCBI Description

(D90905) 50S ribosomal protein L36 [Synechocystis sp.]

Seq. No.

4642 3849 1.R1040 Contig ID 5'-most EST kmv700740822.h1

Seq. No.

4643

Contig ID 3854 1.R1040

5'-most EST uC-gmrominsoy241g01b1

Seq. No.

4644

3854 2.R1040 Contig ID

5'-most EST LIB3052-002-Q1-B1-C2

Method BLASTX a4572675 NCBI GI 203 BLAST score 1.0e-15 E value Match length 146 % identity 42



NCBI Description (AC006954) unknown protein [Arabidopsis thaliana]

Seq. No. 4645

Contig ID 3854 3.R1040

5'-most EST LIB3053-003-Q1-N1-D3

Seq. No. 4646

Contig ID 3855_1.R1040 5'-most EST zhf700956406.h1

Method BLASTX
NCBI GI g3776581
BLAST score 368
E value 5.0e-35
Match length 84
% identity 75

NCBI Description (AC005388) Similar to Beta integral membrane protein

homolog gb_U43629 from A. thaliana. [Arabidopsis thaliana]

Seq. No. 4647

Contig ID 3855 2.R1040

5'-most EST LIB $3\overline{0}$ 28-025-Q1-B1-B5

Method BLASTX
NCBI GI g3776581
BLAST score 187
E value 5.0e-14
Match length 46
% identity 74

NCBI Description (AC005388) Similar to Beta integral membrane protein

homolog gb_U43629 from A. thaliana. [Arabidopsis thaliana]

Seq. No. 4648

Contig ID 3857 1.R1040

5'-most EST uC-qmflminsoy002b11b1

Seq. No. 4649

Contig ID 3857_2.R1040 5'-most EST zhf700963527.h1

Seq. No. 4650

Contig ID 3858_1.R1040

5'-most EST uC-gmropic039f10b1

Method BLASTX
NCBI GI g2809262
BLAST score 698
E value 3.0e-73
Match length 349
% identity 44

NCBI Description (AC002560) F21B7.31 [Arabidopsis thaliana]

Seq. No. 4651

Contig ID 3858 2.R1040

5'-most EST LIB3170-053-Q1-K1-F9

Method BLASTX
NCBI GI g3201632
BLAST score 218
E value 9.0e-28
Match length 150



% identity 45 NCBI Description (AC004669) putative 2A6 protein [Arabidopsis thaliana]

Seq. No. 4652

Contig ID 3858 3.R1040

5'-most EST LIB3170-054-Q1-J1-C3

Method BLASTX
NCBI GI g2809262
BLAST score 308
E value 5.0e-28
Match length 124
% identity 53

NCBI Description (AC002560) F21B7.31 [Arabidopsis thaliana]

Seq. No. 4653

Contig ID 3861 1.R1040 5'-most EST zhf700962573.h1

Seq. No. 4654

Contig ID 3861_2.R1040 5'-most EST zhf700964117.h1

Seq. No. 4655

Contig ID 3861_3.R1040 5'-most EST seb700648626.h1

Seq. No. 4656

Contig ID 3861 4.R1040

5'-most EST LIB3050-011-Q1-E1-H6

Seq. No. 4657

Contig ID 3861_5.R1040 5'-most EST zhf700953349.h1

Method BLASTX
NCBI GI g4454465
BLAST score 193
E value 6.0e-15
Match length 55
% identity 58

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No. 4658

Contig ID 3862 1.R1040

5'-most EST jC-gmst02400044a05a1

Seq. No. 4659

Contig ID 3863_1.R1040

5'-most EST jC-gmro02910007b04a1

Seq. No. 4660

Contig ID 3864 1.R1040

5'-most EST LIB3170-043-Q1-J1-C11

Method BLASTX
NCBI GI g3941522
BLAST score 460
E value 5.0e-46
Match length 100



% identity (AF062915) putative transcription factor [Arabidopsis NCBI Description thaliana] Seq. No. 4661 3868 1.R1040 Contig ID 5'-most EST seb700653512.h1 Method BLASTX NCBI GI q2842486 BLAST score 357 E value 1.0e-33

% identity (AL021749) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 4662

Match length

3868 2.R1040 Contig ID 5'-most EST g4293429

4663 Seq. No.

Contia ID 3868 3.R1040

5'-most EST uC-qmronoir064a11b1

131

4664 Seq. No.

Contig ID 3872 1.R1040 5'-most EST jex700909729.h1

Method BLASTX NCBI GI g1171161 BLAST score 654 1.0e-68 E value 175 Match length % identity 69

(U41472) pectate lyase homolog [Medicago sativa] NCBI Description

4665 Seq. No.

3876 1.R1040 Contig ID jex700904814.h15'-most EST

Method BLASTX g1171579 NCBI GI BLAST score 1857 E value 0.0e + 00Match length 478 73 % identity

(X95342) cytochrome P450 [Nicotiana tabacum] NCBI Description

4666 Seq. No.

3879 1.R1040 Contig ID jex700905393.hl 5'-most EST

Method BLASTX q2895576 NCBI GI BLAST score 557 1.0e-56 E value Match length 433 29 % identity

(AF041337) vacuolar proton pump subunit SFD beta isoform NCBI Description

[Bos taurus]



```
Seq. No.
                   4667
Contig ID
                   3880 1.R1040
5'-most EST
                   leu7\overline{0}1145144.h1
Method
                   BLASTX
                   g886130
NCBI GI
BLAST score
                   443
E value
                   1.0e-43
Match length
                   224
% identity
                   42
NCBI Description
                  (U28148) putative pectinesterase [Medicago sativa]
                   4668
Seq. No.
Contig ID
                   3882 1.R1040
5'-most EST
                   LIB3107-006-Q1-K1-E11
Seq. No.
                   4669
Contig ID
                   3882 2.R1040
5'-most EST
                   LIB3107-031-Q1-K1-H6
Seq. No.
                   4670
Contig ID
                   3888 1.R1040
5'-most EST
                   zpv700759675.h1
Method
                   BLASTX
NCBI GI
                   g1877279
BLAST score
                   367
E value
                   5.0e-35
Match length
                   115
% identity
                   58
NCBI Description (Z92770) fadE2 [Mycobacterium tuberculosis]
Seq. No.
                   4671
Contig ID
                   3889 1.R1040
5'-most EST
                   uC-gmronoir064h11b1
Seq. No.
                   4672
                   3891 1.R1040
Contig ID
5'-most EST
                   bth700847332.h1
Seq. No.
                   4673
Contig ID
                   3897 1.R1040
5'-most EST
                   LIB3028-024-Q1-B1-C7
Seq. No.
                   4674
Contig ID
                   3899 1.R1040
5'-most EST
                   kl1701209291.h1
Method
                   BLASTX
NCBI GI
                   q2244973
BLAST score
                   324
E value
                   1.0e-29
Match length
                   123
% identity
                   58
                   (Z97340) similarity to extensin class 1 protein
NCBI Description
                   [Arabidopsis thaliana]
```

Seq. No.

4675

3900 1.R1040 Contig ID 5'-most EST leu701145246.h1



```
Seq. No.
                   4676
Contig ID
                   3900 2.R1040
5'-most EST
                   LIB3028-021-Q1-B1-G11
Seq. No.
                   4677
Contig ID
                   3900 3.R1040
5'-most EST
                   gsv701045626.h1
Seq. No.
                   4678
Contig ID
                   3902 1.R1040
5'-most EST
                   LIB3028-024-Q1-B1-D12
Method
                   BLASTX
NCBI GI
                   g2960120
BLAST score
                   563
E value
                   1.0e-57
Match length
                   248
% identity
                   47
NCBI Description
                  (AL022121) glpK [Mycobacterium tuberculosis]
Seq. No.
                   4679
Contig ID
                   3904 1.R1040
5'-most EST
                   gsv701051707.hl
Method
                   BLASTX
NCBI GI
                   g2688828
                                                                               Ty.
BLAST score
                   290
E value
                   5.0e-26
Match length
                   121
% identity
                   50
                  (U97530) ethylene-forming-enzyme-like dioxygenase [Prunus
NCBI Description
                   armeniaca]
                   4680
Seq. No.
Contig ID
                   3906 1.R1040
5'-most EST
                  LIB3073-006-Q1-K1-E7
Method
                  BLASTN
NCBI GI
                  g2982267
                  150
BLAST score
E value
                  2.0e-78
Match length
                   422
% identity
                   84
NCBI Description
                  Picea mariana probable 40S ribosomal protein S15 (Sb23)
                  mRNA, complete cds
Seq. No.
                   4681
Contig ID
                   3906 2.R1040
5'-most EST
                  LIB3093-027-Q1-K1-B3
Method
                  BLASTN
NCBI GI
                  g2982267
BLAST score
                  178
                  4.0e-95
E value
Match length
                  422
% identity
                  86
```

Seq. No. 4682

NCBI Description

mRNA, complete cds

Picea mariana probable 40S ribosomal protein S15 (Sb23)

Match length

% identity

162

96



```
Contig ID
                  3906 3.R1040
5'-most EST
                  crh700855117.h1
Method
                  BLASTN
NCBI GI
                  q2982267
BLAST score
                  117
E value
                  8.0e-59
Match length
                  425
% identity
                  83
NCBI Description
                  Picea mariana probable 40S ribosomal protein S15 (Sb23)
                  mRNA, complete cds
Seq. No.
                  4683
Contig ID
                  3910 1.R1040
5'-most EST
                  LIB3049-027-Q1-E1-E3
Method
                  BLASTX
NCBI GI
                  q4539422
BLAST score
                  416
E value
                  9.0e-41
Match length
                  117
% identity
                  68
                  (AL049171) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4684
                  3914 1.R1040
Contig ID
5'-most EST
                  gsv701043825.h1
Method
                  BLASTX
NCBI GI
                  q114200
BLAST score
                  165
E value
                  3.0e-11
Match length
                  117
% identity
                  38
NCBI Description
                  SHIKIMATE KINASE PRECURSOR >gi 100253 pir S21584 shikimate
                  kinase precursor - tomato >qi 19349 emb CAA45121 (X63560)
                  shikimate kinase precursor [Lycopersicon esculentum]
Seq. No.
                  4685
Contig ID
                  3914 2.R1040
5'-most EST
                  g4260282
Method
                  BLASTX
NCBI GI
                  q114200
BLAST score
                  443
E value
                  4.0e-60
Match length
                  256
% identity
                  49
NCBI Description
                  SHIKIMATE KINASE PRECURSOR >gi 100253 pir S21584 shikimate
                  kinase precursor - tomato >gi 19349 emb CAA45121 (X63560)
                  shikimate kinase precursor [Lycopersicon esculentum]
Seq. No.
                  4686
Contig ID
                  3915 1.R1040
5'-most EST
                  LIB3028-024-Q1-B1-B12
Method
                  BLASTX
NCBI GI
                  q3913437
BLAST score
                  822
E value
                  2.0e-88
```



PUTATIVE PRE-MRNA SPLICING FACTOR ATP-DEPENDENT RNA NCBI Description HELICASE >gi_1402875_emb_CAA66825_ (X98130) RNA helicase [Arabidopsis thaliana] >gi_1495271_emb_CAA66613_ (X97970) RNA helicase [Arabidopsis thaliana] Seq. No. 4687 Contig ID 3916 1.R1040 5'-most EST jC-qmro02800031h12d1

4688 Seq. No.

Contig ID 3917 1.R1040

5'-most EST LIB3028-024-Q1-B1-B3

Method BLASTX q4056495 NCBI GI BLAST score 176 E value 1.0e-12 Match length 88 % identity 51

(AC005896) putative TKRP125 [Arabidopsis thaliana] NCBI Description

4689 Seq. No.

Contig ID 3918 1.R1040

5'-most EST LIB3028-024-Q1-B1-B4

Method BLASTX NCBI GI g2829910 BLAST score 1025 E value 1.0e-112 Match length 295 % identity

(AC002291) Unknown protein, contains regulator of NCBI Description

chromosome condensation motifs [Arabidopsis thaliana]

4690 Seq. No.

3921 1.R1040 Contig ID 5'-most EST zzp700833785.h1

Method BLASTX g3881976 NCBI GI BLAST score 530 5.0e-54 E value Match length 155 % identity 63

(AJ012409) hypothetical protein [Homo sapiens] NCBI Description

4691 Seq. No.

3922 1.R1040 Contig ID

LIB3049-042-Q1-E1-H3 5'-most EST

Method BLASTX g1353516 NCBI GI BLAST score 2150 0.0e + 00E value Match length 501 % identity

NCBI Description (U38651) sugar transporter [Medicago truncatula]

Seq. No. 4692

3922 2.R1040 Contig ID g4282909 5'-most EST



Method BLASTN
NCBI GI g169717
BLAST score 54
E value 2.0e-21
Match length 110
% identity 87

NCBI Description Ricinus communis (clone PST293) sugar carrier protein

(RCSTC) mRNA, complete CDS

Seq. No. 4693

Contig ID 3922 3.R1040

5'-most EST uC-gmrominsoy258d01b1

Method BLASTX
NCBI GI g3915039
BLAST score 270
E value 2.0e-23
Match length 60
% identity 90

NCBI Description SUGAR CARRIER PROTEIN C >gi 169718 (L08196) sugar carrier

protein [Ricinus communis]

Seq. No. 4694

Contig ID 3922 4.R1040

5'-most EST fC-gmse7000755922r1

Method BLASTX
NCBI GI g1083942
BLAST score 651
E value 6.0e-68
Match length 254
% identity 48

NCBI Description rubber particle cytochrome P450 - guayule

>gi 791093 emb CAA55025 (X78166) rubber particle protein

[Parthenium argentatum]

Seq. No. 4695

Contig ID 3922 5.R1040

5'-most EST uC-gmrominsoy140g06b1

Method BLASTX
NCBI GI g1353516
BLAST score 431
E value 2.0e-42
Match length 120
% identity 72

NCBI Description (U38651) sugar transporter [Medicago truncatula]

Seq. No. 4696

Contig ID 3922_6.R1040 5'-most EST zhf700959335.h1

Seq. No. 4697

Contig ID 3922 8.R1040

5'-most EST uC-qmronoir024d11b1

Method BLASTN
NCBI GI g1353515
BLAST score 63
E value 5.0e-27
Match length 123



```
% identity
NCBI Description Medicago truncatula sugar tranporter mRNA, complete cds
Seq. No.
                  4698
                  3922 10.R1040
Contig ID
                  vwf700674586.hl
5'-most EST
                  4699
Seq. No.
                  3923 1.R1040
Contig ID
                  LIB3028-024-Q1-B1-C10
5'-most EST
                  4700
Seq. No.
                  3925 1.R1040
Contig ID
                  LIB3050-001-Q1-E1-F4
5'-most EST
Method
                  BLASTX
                  q3759184
NCBI GI
                  1136
BLAST score
                  1.0e-124
E value
                  302
Match length
                  73
% identity
NCBI Description (AB018441) phi-1 [Nicotiana tabacum]
                   4701
Seq. No.
                  3925 2.R1040
Contig ID
5'-most EST
                  LIB3\overline{1}07-004-Q1-K1-F3
Method
                  BLASTX
                   g3759184
NCBI GI
BLAST score
                  728
                   4.0e-77
E value
                   222
Match length
% identity
                   68
                  (AB018441) phi-1 [Nicotiana tabacum]
NCBI Description
                   4702
Seq. No.
                   3928 1.R1040
Contig ID
5'-most EST
                   leu701156935.h1
Method
                   BLASTX
                   q3413322
NCBI GI
BLAST score
                   688
E value
                   2.0e-72
                   191
Match length
% identity
NCBI Description (Y11118) polygalacturonase [Medicago sativa]
Seq. No.
                   4703
                   3929 1.R1040
Contig ID
5'-most EST
                   LIB3028-023-Q1-B1-H5
```

Seq. No.

4704

Contig ID 3932 1.R1040

5'-most EST LIB3092-017-Q1-K1-A3

Method BLASTX q1076288 NCBI GI 393 BLAST score 7.0e-38 E value Match length 149 % identity 48



NCBI Description amino acid permease AAP3 - Arabidopsis thaliana

 Seq. No.
 4705

 Contig ID
 3937_1.R1040

 5'-most EST
 hyd700730315.h1

 Method
 BLASTN

 NCBI GI
 g20872

 BLAST score
 217

E value 1.0e-118 Match length 525 % identity 85

NCBI Description Pea mRNA for plastid ribosomal protein CL24

Seq. No. 4706

Contig ID 3940 1.R1040

5'-most EST LIB3170-042-Q1-K1-A6

Method BLASTX
NCBI GI g285741
BLAST score 1512
E value 1.0e-168
Match length 422
% identity 70

NCBI Description (D14550) EDGP precursor [Daucus carota]

Seq. No. 4707

Contig ID 3941_1.R1040 5'-most EST fua701042582.h1

Method BLASTN
NCBI GI g1620908
BLAST score 239
E value 1.0e-131
Match length 523
% identity 86

NCBI Description Carrot mRNA for DcARF1, complete cds

Seq. No. 4708

Contig ID 3941 2.R1040

5'-most EST LIB3109-035-Q1-K1-D6

Method BLASTN
NCBI GI g1620908
BLAST score 235
E value 1.0e-129
Match length 531
% identity 86

NCBI Description Carrot mRNA for DcARF1, complete cds

Seq. No.

4709

Contig ID 3941 3.R1040

5'-most EST LIB3106-060-Q1-K1-G9

Method BLASTX
NCBI GI g2281102
BLAST score 213
E value 1.0e-16
Match length 93
% identity 71

NCBI Description (AC002333) SF16 isolog [Arabidopsis thaliana]

S. .



```
4710
Seq. No.
                  3941 4.R1040
Contig ID
5'-most EST
                  jex700910011.hl
Method
                  BLASTX
NCBI GI
                  g3819099
BLAST score
                  211
                  6.0e-17
E value
Match length
                  75
% identity
                  56
NCBI Description (AJ009825) copper amine oxidase [Cicer arietinum]
Seq. No.
                  4711
                  3941 5.R1040
Contig ID
5'-most EST
                  LIB3109-035-Q1-K6-D12
Method
                  BLASTX
NCBI GI
                  g2501337
BLAST score
                  432
E value
                  6.0e-46
Match length
                  118
% identity
                  75
                  AMINE OXIDASE PRECURSOR [COPPER-CONTAINING]
NCBI Description
                  >gi 2129875 pir C44239 amine oxidase (copper-containing)
                  (EC 1.4.3.6) precursor - garden pea >gi_685198 (L39931)
                  copper amine oxidase [Pisum sativum]
Seq. No.
                  4712
Contig ID
                  3941 6.R1040
5'-most EST
                  kl1701214261.h1
                  4713
Seq. No.
                  3941 7.R1040
Contig ID
5'-most EST
                  LIB3074-011-Q1-E1-B4
Seq. No.
                  4714
                  3941_9.R1040
Contig ID
5'-most EST
                  jC-gmst02400039f10a1
Method
                  BLASTN
NCBI GI
                  g1620908
BLAST score
                  42
                  3.0e-14
E value
Match length
                  70
% identity
                  90
NCBI Description Carrot mRNA for DcARF1, complete cds
Seq. No.
                  4715
Contig ID
                  3941 10.R1040
                  jC-gmst02400047a03d1
5'-most EST
                  BLASTN
Method
```

NCBI GI BLAST score g1620908

E value

59 2.0e-24

Match length % identity

135 86

NCBI Description Carrot mRNA for DcARF1, complete cds

Seq. No.

4716

Contig ID

3941 12.R1040



```
LIB3087-012-Q1-K1-A5
5'-most EST
                  BLASTN
Method
                  g1184986
NCBI GI
BLAST score
                  76
                  2.0e-34
E value
                  160
Match length
                  87
% identity
                  Nicotiana tabacum GTP-binding protein NTGB1 mRNA, partial
NCBI Description
                   4717
Seq. No.
                  3942 1.R1040
Contig ID
                  LIB3040-008-Q1-E1-G9
5'-most EST
                  BLASTN
Method
NCBI GI
                  g170919
BLAST score
                  122
                  8.0e-62
E value
                  322
Match length
% identity
                  85
                  C.maltosa ribosomal protein L41 (LEL41) gene, complete cds
NCBI Description
                   4718
Seq. No.
                   3943 1.R1040
Contig ID
                  pmv700895034.h1
5'-most EST
Seq. No.
                   4719
                   3944 1.R1040
Contig ID
5'-most EST
                  pmv700893947.h1
Method
                   BLASTX
                   q3894194
NCBI GI
                   479
BLAST score
                   7.0e-48
E value
Match length
                   158
                   56
% identity
                   (AC005662) putative strictosidine synthase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   4720
Contig ID
                   3945 1.R1040
5'-most EST
                  LIB3039-042-Q1-E1-C11
                   4721
Seq. No.
                   3945 2.R1040
Contig ID
5'-most EST
                   smc7\overline{0}0747713.h1
                   4722
Seq. No.
                   3946 1.R1040
Contig ID
                   LIB3049-002-Q1-E1-D7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3283893
BLAST score
                   143
E value
                   8.0e-12
```

Match length 87 % identity 43

NCBI Description (AF070626) unknown [Homo sapiens]

Seq. No. 4723



```
3946 2.R1040
Contig ID
                  LIB3170-029-Q1-K1-D3
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3283893
BLAST score
                  157
                  2.0e-10
E value
                  70
Match length
% identity
                  46
NCBI Description (AF070626) unknown [Homo sapiens]
Seq. No.
                  4724
                  3946 3.R1040
Contig ID
                  jC-gmf102220063h02a1
5'-most EST
Seq. No.
                  4725
Contig ID
                  3946 4.R1040
                  LIB3073-006-Q1-K1-A12
5'-most EST
Seq. No.
                  4726
Contig ID
                  3947 1.R1040
5'-most EST
                  LIB3028-023-Q1-B1-G10
Method
                  BLASTX
NCBI GI
                  g2832640
BLAST score
                  592
E value
                  2.0e-61
Match length
                  159
% identity
                  69
                  (AL021710) neoxanthin cleavage enzyme - like protein
NCBI Description
                  [Arabidopsis thaliana]
Seq. No.
                  4727
                  3949 1.R1040
Contig ID
                  fua701038069.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g1174583
                  298
BLAST score
                  1.0e-26
E value
Match length
                  162
% identity
                  40
                  TRANSALDOLASE >gi 1074653 pir D64167 hypothetical protein
NCBI Description
                  HI1125 - Haemophilus influenzae (strain Rd KW20)
                  >gi 1574680 (U32792) transaldolase B (talB) [Haemophilus
                  influenzae Rd]
                  4728
Seq. No.
                  3949 2.R1040
Contig ID
5'-most EST
                  asn701140690.h1
Seq. No.
                   4729
                  3951 1.R1040
Contig ID
5'-most EST
                  uC-qmrominsoy048d09b1
```

Method BLASTX
NCBI GI g2499710
BLAST score 1620
E value 0.0e+00
Match length 346
% identity 86



NCBI Description PHOSPHOLIPASE D PRECURSOR (PLD) (CHOLINE PHOSPHATASE) (PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D) >gi_1438075 (L33686) phospholipase D [Ricinus communis]

4730 Seq. No.

Contig ID 3951 2.R1040

LIB3051-105-Q1-K1-G10 5'-most EST

Method BLASTX q3914359 NCBI GI 1877 BLAST score 0.0e+00E value 382 Match length % identity

PHOSPHOLIPASE D PRECURSOR (PLD) (CHOLINE PHOSPHATASE) NCBI Description

(PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D)

>gi_1928979 (U92656) phospholipase D [Vigna unguiculata]

Seq. No. 4731

3952 1.R1040 Contig ID asn701138149.hl 5'-most EST

BLASTX Method NCBI GI a3024706 280 BLAST score 2.0e-24 E value 94 Match length

% identity

TRANSCRIPTION INITIATION FACTOR TFIID 18 KD SUBUNIT NCBI Description

(TAFII-18) (TAFII18) >gi 1362894_pir__S54782 PolII transcription factor TFIID chain hTAFII18 - human >gi_791053_emb_CAA58827_ (X84003) PolII transcription

factor TFIID [Homo sapiens]

4732 Seq. No.

3952 2.R1040 Contig ID

LIB3051-034-Q1-K1-C10 5'-most EST

4733 Seq. No.

3952 3.R1040 Contig ID

LIB3028-023-Q1-B1-G5 5'-most EST

Seq. No.

4734

3953 1.R1040 Contig ID

LIB3028-023-Q1-B1-G6 5'-most EST

Seq. No.

4735

Contig ID 3956 1.R1040 $zhf7\overline{0}0956881.h1$ 5'-most EST

BLASTX Method NCBI GI q2494299 BLAST score 1032 1.0e-112 E value Match length 269 % identity 76

EUKARYOTIC TRANSLATION INITIATION FACTOR 2 BETA SUBUNIT NCBI Description

(EIF-2-BETA) >gi 1732361 (U80269) translation initiation

factor 2 beta [Malus domestica]



```
4736
Seq. No.
                  3957 1.R1040
Contig ID
5'-most EST
                  LIB3139-055-P1-N1-E11
Method
                  BLASTN
                  g408793
NCBI GI
BLAST score
                  698
                  0.0e+00
E value
Match length
                  801
% identity
                  98
                  Glycine soja chloroplast 3-omega faty acid desaturase
NCBI Description
                  (Fad3) mRNA, complete cds
                  4737
Seq. No.
Contig ID
                  3957 2.R1040
5'-most EST
                  uC-gmflminsoy036a11b1
Method
                  BLASTN
                  g408793
NCBI GI
BLAST score
                  235
                  1.0e-129
E value
Match length
                  635
% identity
                  92
NCBI Description Glycine soja chloroplast 3-omega faty acid desaturase
                  (Fad3) mRNA, complete cds
Seq. No.
                  4738
Contig ID
                  3958 1.R1040
5'-most EST
                  LIB3028-023-Q1-B1-D4
                  4739
Seq. No.
Contig ID
                  3961 1.R1040
5'-most EST
                  zsg701120931.h1
Method
                  BLASTX
                  q3184061
NCBI GI
BLAST score
                   340
                  1.0e-31
E value
Match length
                  177
% identity
NCBI Description
                  (AL023776) atp dependent helicase [Schizosaccharomyces
                  pombe]
                   4740
Seq. No.
Contig ID
                   3962 1.R1040
5'-most EST
                  LIB3049-025-Q1-E1-H8
Seq. No.
                   4741
                   3964 1.R1040
Contig ID
5'-most EST
                   jC-gmst02400020dg12d1
Seq. No.
                   4742
                   3968 1.R1040
Contig ID
```

5'-most EST

LIB3028-023-Q1-B1-E4

Seq. No. Contig ID 4743 3970 1.R1040

5'-most EST

jC-gmro02800029h10a1

Method

BLASTX

NCBI GI

g3329368



```
735
BLAST score
                   1.0e-77
E value
Match length
                   377
% identity
                   43
                  (AF031244) nodulin-like protein [Arabidopsis thaliana]
NCBI Description
                   4744
Seq. No.
                   3971 1.R1040
Contig ID
5'-most EST
                   dpv7\overline{0}1102529.h1
Method
                   BLASTX
                   g481896
NCBI GI
BLAST score
                   498
                   8.0e-50
E value
Match length
                   246
% identity
                   41
NCBI Description
                  chitinase (EC 3.2.1.14) - rice
                   4745
Seq. No.
                   3971 2.R1040
Contig ID
                   jC-gmf102220085d04d1
5'-most EST
Seq. No.
                   4746
                   3971 3.R1040
Contig ID
                   pxt700945993.h1
5'-most EST
Seq. No.
                   4747
                   3972 1.R1040
Contig ID
5'-most EST
                   bth700847151.hl
Method
                   BLASTX
                   q2829751
NCBI GI
BLAST score
                   205
                   8.0e-16
E value
                   114
Match length
% identity
                   37
                   MACROPHAGE MIGRATION INHIBITORY FACTOR HOMOLOG (BMMIF)
NCBI Description
                   >gi 1850559 (U88035) macrophage migration inhibitory factor
                   [Brugia malayi] >gi_2190976 (AF002699) macrophage migration
                   inhibitory factor [Brugia malayi]
                   4748
Seq. No.
Contig ID
                   3974 1.R1040
5'-most EST
                   LIB3028-023-Q1-B1-B7
                   4749
Seq. No.
                   3979 1.R1040
Contig ID
                   LIB3028-023-Q1-B1-C3
5'-most EST
Method
                   BLASTX
                   q3702333
NCBI GI
BLAST score
                   196
E value
                   4.0e-15
Match length
                   95
```

NCBI Description

% identity

49

(AC005397) hypothetical protein [Arabidopsis thaliana]



```
Method
                  BLASTX
NCBI GI
                  q4406772
BLAST score
                  306
E value
                  2.0e-27
Match length
                  119
% identity
                   55
                   (AC006836) putative nitrilase-associated protein
NCBI Description
                   [Arabidopsis thaliana]
                   4751
Seq. No.
Contig ID
                  3983 2.R1040
5'-most EST
                  sat701003125.h1
Method
                  BLASTX
NCBI GI
                  g2765837
BLAST score
                  279
E value
                   2.0e-24
Match length
                  131
% identity
                   50
                  (Z96936) NAP16kDa protein [Arabidopsis thaliana]
NCBI Description
                   4752
Seq. No.
Contia ID
                   3983 3.R1040
5'-most EST
                  jC-gmro02910061a08a1
Seq. No.
                   4753
Contig ID
                   3983 4.R1040
5'-most EST
                  LIB3053-006-Q1-N1-A8
Method
                   BLASTX
NCBI GI
                   g2765837
BLAST score
                   208
                   4.0e-16
E value
Match length
                   102
                   49
% identity
                  (Z96936) NAP16kDa protein [Arabidopsis thaliana]
NCBI Description
                   4754
Seq. No.
                   3987 1.R1040
Contig ID
5'-most EST
                   pmv700894858.hl
Method
                   BLASTX
NCBI GI
                   g3860246
BLAST score
                   172
E value
                   1.0e-11
                   153
Match length
% identity
                   36
                   (AC005824) putative reverse-transcriptase protein
NCBI Description
                   [Arabidopsis thaliana] >gi_4510429_gb_AAD21515.1
                   (AC006929) putative reverse transcriptase [Arabidopsis
                   thaliana]
Seq. No.
                   4755
Contig ID
                   3988 1.R1040
                   ssr700556516.hl
5'-most EST
```

Method

BLASTX

NCBI GI g2245012 BLAST score 172 E value 4.0e-21 Match length 91

BLAST score

E value

461

1.0e-46



```
% identity
NCBI Description
                  (Z97341) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  4756
                  3990 1.R1040
Contig ID
5'-most EST
                  uC-gmropic018e11b1
Method
                  BLASTX
NCBI GI
                  q4107276
BLAST score
                  1010
E value
                  1.0e-137
Match length
                  271
% identity
                  85
                  (X98506) acetyl-CoA synthetase [Solanum tuberosum]
NCBI Description
Seq. No.
                   4757
Contig ID
                  3991 1.R1040
5'-most EST
                  g4287582
Method
                  BLASTX
NCBI GI
                  g4335745
BLAST score
                   387
E value
                   4.0e-37
Match length
                  192
% identity
                   42
                   (AC006284) putative hydrolase (contains an
NCBI Description
                   esterase/lipase/thioesterase active site serine domain
                   (prosite: PS50187) [Arabidopsis thaliana]
Seq. No.
                   4758
Contig ID
                   3992 1.R1040
5'-most EST
                  uC-gmrominsoy216h02b1
Method
                  BLASTX
                  q2196466
NCBI GI
BLAST score
                  256
                  2.0e-21
E value
Match length
                  267
% identity
                   44
                  (Y13673) TATA binding protein-associated factor
NCBI Description
                   [Arabidopsis thaliana]
                   4759
Seq. No.
Contig ID
                   3993 1.R1040
5'-most EST
                  LIB3028-023-Q1-B1-A2
Method
                  BLASTX
NCBI GI
                   g3510248
BLAST score
                   200
                   2.0e-15
E value
Match length
                   94
                   50
% identity
                  (AC005310) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4760
Contig ID
                   4001 1.R1040
5'-most EST
                   txt700732386.h1
Method
                  BLASTX
NCBI GI
                  g2911044
```

920



```
225
Match length
% identity
                   49
                   (AL021961) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4761
                   4001 2.R1040
Contig ID
5'-most EST
                  LIB3170-074-Q1-J1-A6
                   4762
Seq. No.
Contig ID
                   4007 1.R1040
```

LIB3028-022-Q1-B1-G10 5'-most EST Method BLASTX g3924611 NCBI GI BLAST score 584 E value 1.0e-70 Match length 207

67 % identity

Seq. No.

(AF069442) hypothetical protein [Arabidopsis thaliana] NCBI Description

4011 1.R1040 Contig ID LIB3073-024-Q1-K1-F6 5'-most EST Method BLASTN g166421 NCBI GI BLAST score 163

E value 3.0e-86 Match length 459 % identity 84

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

4763

4764 Seq. No.

4011 2.R1040 Contig ID

uC-gmronoir062f10b1 5'-most EST

Method BLASTN q166421 NCBI GI BLAST score 171 E value 6.0e-91 459 Match length 84 % identity

Medicago sativa ubiquitin carrier protein mRNA, complete NCBI Description

4765 Seq. No.

4011 3.R1040 Contig ID xpa700794105.h1 5'-most EST

BLASTN Method NCBI GI q166923 BLAST score 78 9.0e-36 E value Match length 234 % identity

Arabidopsis thaliana ubiquitin carrier protein (UBC1) mRNA, NCBI Description

complete cds

Seq. No. 4766

4011 6.R1040 Contig ID

NCBI Description

tabacum]



```
g5510194
5'-most EST
                   BLASTN
Method
                   g166421
NCBI GI
BLAST score
                   38
                   5.0e-12
E value
                   86
Match length
% identity
                   86
                   Medicago sativa ubiquitin carrier protein mRNA, complete
NCBI Description
                   4767
Seq. No.
                   4015 1.R1040
Contig ID
                   k117\overline{0}1211906.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3004556
BLAST score
                   363
                   2.0e-34
E value
                   130
Match length
% identity
                   67
                   (AC003673) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   4768
Seq. No.
                   4019 1.R1040
Contig ID
                   jC-gmro02910021b09d1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3021506
BLAST score
                   1639
E value
                   0.0e + 00
Match length
                   359
                   89
% identity
                   (X96727) isocitrate dehydrogenase (NAD+) [Nicotiana
NCBI Description
                   tabacum]
                   4769
Seq. No.
                   4019 2.R1040
Contig ID
5'-most EST
                   LIB3051-090-Q1-K1-G6
                   BLASTX
Method
NCBI GI
                   g3021506
BLAST score
                   313
E value
                   1.0e-28
Match length
                   90
                   72
% identity
                   (X96727) isocitrate dehydrogenase (NAD+) [Nicotiana
NCBI Description
                   tabacum]
                   4770
Seq. No.
                   4019 3.R1040
Contig ID
                   LIB3093-007-Q1-K1-H10
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3021506
BLAST score
                   273
                   4.0e-24
E value
Match length
                   80
% identity
                   72
```

(X96727) isocitrate dehydrogenase (NAD+) [Nicotiana



```
4771
Seq. No.
                  4019 5.R1040
Contig ID
5'-most EST
                  jex700904639.hl
                   4772
Seq. No.
                  4020 1.R1040
Contig ID
5'-most EST
                  LIB3039-027-Q1-E1-B10
                  BLASTX
Method
NCBI GI
                  g2695711
BLAST score
                  619
                  3.0e-64
E value
                  134
Match length
                  84
% identity
                  (AJ001370) cytochome b5 [Olea europaea]
NCBI Description
                  4773
Seq. No.
                   4020 2.R1040
Contig ID
                   fC-qmro700848840a1
5'-most EST
Method
                  BLASTN
NCBI GI
                   g2647950
                   59
BLAST score
                  2.0e-24
E value
                  123
Match length
                  87
% identity
NCBI Description Olea europaea cytochrome b5 gene-2
                   4774
Seq. No.
                   4020 3.R1040
Contig ID
                  bth700848441.h1
5'-most EST
                  BLASTX
Method
                   g4240120
NCBI GI
BLAST score
                   226
                   9.0e-19
E value
Match length
                   64
% identity
                   69
NCBI Description
                  (AB007801) cytochrome b5 [Arabidopsis thaliana]
Seq. No.
                   4775
                   4024 1.R1040
Contig ID
5'-most EST
                   LIB3028-022-Q1-B1-E10
                   BLASTX
Method
NCBI GI
                   q4558665
BLAST score
                   170
E value
                   5.0e-12
Match length
                   124
                   29
% identity
NCBI Description
                  (AC007063) putative white protein [Arabidopsis thaliana]
Seq. No.
                   4776
Contig ID
                   4026 1.R1040
                   hyd700725705.h1
5'-most EST
```

5'-most EST hyd700725'
Method BLASTX
NCBI GI g2462753
BLAST score 306
E value 4.0e-32
Match length 129

55

% identity



```
(AC002292) putative polygalacturonase [Arabidopsis
NCBI Description
                  thaliana]
                  4777
Seq. No.
Contig ID
                  4028 1.R1040
5'-most EST
                  jex700908404.hl
                  BLASTX
Method
NCBI GI
                  q4218120
BLAST score
                  283
                  4.0e-25
E value
Match length
                  90
% identity
                  64
                  (AL035353) Proline-rich APG-like protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  4778
Contig ID
                  4028 2.R1040
5'-most EST
                  LIB3030-004-Q1-B1-G8
                  BLASTX
Method
                  q4218120
NCBI GI
BLAST score
                  444
E value
                  4.0e-44
                  109
Match length
                  78
% identity
                  (AL035353) Proline-rich APG-like protein [Arabidopsis
NCBI Description
                  thaliana]
                   4779
Seq. No.
Contig ID
                   4031 1.R1040
5'-most EST
                  wrg700788672.h1
Method
                  BLASTX
                  g3702327
NCBI GI
BLAST score
                  576
                  3.0e-59
E value
                  162
Match length
                  75
% identity
                  (AC005397) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4780
Seq. No.
Contig ID
                   4031 3.R1040
                  LIB3056-010-Q1-N1-H8
5'-most EST
                   4781
Seq. No.
                   4037 1.R1040
Contig ID
                  LIB3049-042-Q1-E1-D9
5'-most EST
                  BLASTX
Method
                   g3342800
NCBI GI
BLAST score
                   848
                   4.0e-91
E value
Match length
                   211
```

Seq. No. 4782

% identity

NCBI Description

Contig ID 4037 2.R1040

5'-most EST LIB3028-022-Q1-B1-D5

78

dehydrogenase [Zea mays]

(AF061837) putative cytosolic 6-phosphogluconate



```
BLASTX
Method
                   q3342802
NCBI GI
                   461
BLAST score
                   1.0e-45
E value
                   133
Match length
% identity
                   (AF061838) putative cytosolic 6-phosphogluconate
NCBI Description
                   dehydrogenase [Zea mays]
                   4783
Seq. No.
                   4037 3.R1040
Contig ID
                   fua701038701.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3342802
                   261
BLAST score
                   1.0e-22
E value
Match length
                   69
                   75
% identity
                   (AF061838) putative cytosolic 6-phosphogluconate
NCBI Description
                   dehydrogenase [Zea mays]
                   4784
Seq. No.
                   4037 4.R1040
Contig ID
                   uC-gmropic063f06b1
5'-most EST
                   BLASTX
Method
                   q2529229
NCBI GI
BLAST score
                   751
                   6.0e-80
E value
                   195
Match length
                   74
% identity
                   (AB007907) 6-phosphogluconate dehydrogenase [Glycine max]
NCBI Description
Seq. No.
                   4785
                   4038 1.R1040
Contig ID
                   LIB3139-014-P1-N1-E12
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1431629
BLAST score
                   1863
                   0.0e+00
E value
                   399
Match length
% identity
                   85
                   (X99348) pectinacetylesterase precursor [Vigna radiata]
NCBI Description
                   4786
Seq. No.
                   4038 2.R1040
Contig ID
                   uC-gmflminsoy119e06b1
 5'-most EST
                   BLASTX
Method
                   g1431629
NCBI GI
                    754
BLAST score
                   3.0e-80
 E value
                   203
Match length
 % identity
                    67
                   (X99348) pectinacetylesterase precursor [Vigna radiata]
 NCBI Description
                    4787
 Seq. No.
```

925

4038 3.R1040

dpv701102573.hl

Contig ID

5'-most EST



Method BLASTN
NCBI GI g1431628
BLAST score 248
E value 1.0e-137
Match length 356
% identity 92

NCBI Description V.radiata mRNA for pectinacetylesterase

Seq. No. 4788

Contig ID 4038 4.R1040

5'-most EST jC-gmro02910008e06a1

Method BLASTN
NCBI GI g1431628
BLAST score 89
E value 3.0e-42
Match length 196
% identity 90

NCBI Description V.radiata mRNA for pectinacetylesterase

Seq. No. 4789

Contig ID 4040 1.R1040

5'-most EST LIB3109-021-Q1-K1-B11

Method BLASTX
NCBI GI g1170508
BLAST score 1957
E value 0.0e+00
Match length 399
% identity 94

NCBI Description EUKARYOTIC INITIATION FACTOR 4A-8 (EIF-4A-8)

>gi_2119931_pir__S60244 translation initiation factor

eIF-4A.8, anther-specific - common tobacco

 $>gi_475219_{emb}_{CAA55639}_{(X79004)}$ translation initiation

factor (eIF-4A) [Nicotiana tabacum]

>gi_475221_emb_CAA55640_ (X79005) translation initiation

factor (eIF-4A) [Nicotiana tabacum]

Seq. No. 4790

4040 2.R1040 Contig ID 5'-most EST g5606855 Method BLASTX g1170508 NCBI GI 974 BLAST score 1.0e-106 E value Match length 202 92 % identity

NCBI Description EUKARYOTIC INITIATION FACTOR 4A-8 (EIF-4A-8)

>gi_2119931_pir__S60244 translation initiation factor

eIF-4A.8, anther-specific - common tobacco

>gi_475219_emb_CAA55639_ (X79004) translation initiation

factor (eIF-4A) [Nicotiana tabacum]

>gi_475221_emb_CAA55640_ (X79005) translation initiation

factor (eIF-4A) [Nicotiana tabacum]

Seq. No. 4791

Contig ID 4040_3.R1040 5'-most EST ncj700986770.h1

Method BLASTX



g2119933 NCBI GI BLAST score 602 2.0e-62 E value Match length 125 % identity 92 NCBI Description translation initiation factor eIF-4A.11 - common tobacco 4792 Seq. No. 4040 4.R1040 Contig ID jC-gmst02400031b10d1 5'-most EST Method BLASTX NCBI GI g1170508 BLAST score 424 1.0e-41 E value 92 Match length % identity EUKARYOTIC INITIATION FACTOR 4A-8 (EIF-4A-8) NCBI Description >gi_2119931_pir__S60244 translation initiation factor

eIF-4A.8, anther-specific - common tobacco

>qi 475219 emb CAA55639 (X79004) translation initiation

factor (eIF-4A) [Nicotiana tabacum]

>gi_475221_emb_CAA55640_ (X79005) translation initiation

factor (eIF-4A) [Nicotiana tabacum]

4793 Seq. No.

Contig ID 4040 5.R1040

LIB3107-067-Q1-K1-C4 5'-most EST

Method BLASTX g2119933 NCBI GI BLAST score 456 2.0e-45 E value Match length 94 93 % identity

NCBI Description translation initiation factor eIF-4A.11 - common tobacco

4794 Seq. No.

4046 1.R1040 Contig ID $k117\overline{0}1206255.h1$ 5'-most EST

BLASTX Method g4406801 NCBI GI 309 BLAST score 6.0e-29 E value 121 Match length 52 % identity

NCBI Description (AC006304) unknown protein [Arabidopsis thaliana]

4795 Seq. No.

4047_1.R1040 Contig ID 5'-most EST crh700851546.h1

Method BLASTN NCBI GI q22635 349 BLAST score 0.0e + 00E value 617 Match length 90 % identity

NCBI Description P.vulgaris mRNA for 70 kD heat shock protein



```
4796
Seq. No.
                  4047 2.R1040
Contig ID
                  jC-gmf102220069b04d1
5'-most EST
Method
                  BLASTN
NCBI GI
                  g22635
BLAST score
                  73
                  7.0e-33
E value
                  160
Match length
% identity
                  87
                  P.vulgaris mRNA for 70 kD heat shock protein
NCBI Description
                  4797
Seq. No.
                  4050 1.R1040
Contig ID
                  wvk700683438.hl
5'-most EST
                  4798
Seq. No.
                  4051 1.R1040
Contig ID
                  LIB3028-022-Q1-B1-C6
5'-most EST
                  BLASTX
Method
                  q4388728
NCBI GI
BLAST score
                  173
                  1.0e-13
E value
                  84
Match length
                  58
% identity
                   (AC006413) putative grr1-like protein [Arabidopsis
NCBI Description
                  thaliana]
                   4799
Seq. No.
                   4053 1.R1040
Contig ID
                  LIB3030-003-Q1-B1-E5
5'-most EST
                  BLASTX
Method
                   g1199467
NCBI GI
BLAST score
                   707
E value
                   8.0e-75
Match length
                   163
                   70
% identity
                   (D64155) possible aldehyde decarbonylase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   4800
                   4054 1.R1040
Contig ID
5'-most EST
                   fC-qmle700683850f1
Method
                   BLASTX
NCBI GI
                   g2352812
BLAST score
                   516
E value
                   3.0e-52
Match length
                   179
% identity
                   56
                   (AF008597) desacetoxyvindoline-4-hydroxylase [Catharanthus
NCBI Description
                   roseus]
                   4801
Seq. No.
```

Contig ID 4055_1.R1040 5'-most EST hrw701058906.h1

Method BLASTN NCBI GI g2224732 BLAST score 109



```
E value 3.0e-54
Match length 356
% identity 87
NCBI Description Vigna radiata mRNA for Aux22e, complete cds
Seg. No. 4802
```

Seq. No. 4802 Contig ID 4055_2.R1040 5'-most EST fjg700968470.h1

Method BLASTN
NCBI GI g2224732
BLAST score 112
E value 5.0e-56
Match length 331
% identity 89

NCBI Description Vigna radiata mRNA for Aux22e, complete cds

Seq. No. 4803

Contig ID 4056_1.R1040 5'-most EST bth700849157.h1

Seq. No. 4804

Contig ID 4056_2.R1040 5'-most EST epx701108971.h1

Seq. No. 4805

Contig ID 4061_1.R1040

5'-most EST LIB3109-011-Q1-K1-H4

Method BLASTX
NCBI GI g2443886
BLAST score 147
E value 5.0e-09
Match length 136
% identity 33

% identity 33

NCBI Description (AC002294) Unknown protein [Arabidopsis thaliana]

Seq. No. 4806

Contig ID 4063_1.R1040

5'-most EST LIB3049-042-Q1-E1-G12

Seq. No. 4807

Contig ID 4063 2.R1040

5'-most EST LIB $3\overline{0}49$ -006-Q1-E1-F2

Seq. No. 4808

Contig ID 4064_1.R1040

5'-most EST LIB3139-016-P1-N1-C10

Method BLASTX
NCBI GI g2924781
BLAST score 509
E value 7.0e-52
Match length 126
% identity 76

NCBI Description (AC002334) putative cellulose synthase [Arabidopsis

thaliana]

Seq. No. 4809

Contig ID 4066 1.R1040



```
pmv700890148.h1
5'-most EST
                  BLASTX
Method
                  g1747310
NCBI GI
                  849
BLAST score
                  5.0e-91
E value
                  314
Match length
                  54
% identity
                  (D58424) Myb-like DNA binding protein [Arabidopsis
NCBI Description
                  thaliana]
                  4810
Seq. No.
                  4067 1.R1040
Contig ID
                  txt700732591.h1
5'-most EST
                  BLASTX
Method
                  q4454452
NCBI GI
                  313
BLAST score
                  1.0e-28
E value
                  100
Match length
                  62
% identity
                   (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4811
Seq. No.
                   4068 1.R1040
Contig ID
                  LIB3107-006-Q1-K1-G1
5'-most EST
                  BLASTX
Method
                   g1781299
NCBI GI
                   289
BLAST score
                   1.0e-25
E value
                   82
Match length
                   66
% identity
                   (Y09506) transformer-SR ribonucleoprotein [Nicotiana
NCBI Description
                   tabacum]
                   4812
Seq. No.
                   4068 2.R1040
Contig ID
5'-most EST
                   LIB3170-054-Q1-K1-A1
Seq. No.
                   4813
Contig ID
                   4069 1.R1040
5'-most EST
                   fC-qmle700685876f1
                   BLASTX
Method
                   q4432858
NCBI GI
BLAST score
                   317
E value
                   6.0e-29
Match length
                   231
% identity
                   (AC006300) putative disease resistance protein [Arabidopsis
NCBI Description
                   thaliana]
                   4814
Seq. No.
                   4074 1.R1040
Contig ID
                   LIB3138-045-Q1-N1-F6
5'-most EST
                   BLASTX
Method
                   g4512678
```

Method BLASTX
NCBI GI g4512678
BLAST score 947
E value 1.0e-102
Match length 263



```
% identity
                  (AC006931) unknown protein [Arabidopsis thaliana]
NCBI Description
                  4815
Seq. No.
Contig ID
                  4074 2.R1040
5'-most EST
                  LIB3138-074-P1-N1-H5
                  BLASTX
Method
                  q4512678
NCBI GI
BLAST score
                  474
                   2.0e-47
E value
Match length
                  158
                   61
% identity
                  (AC006931) unknown protein [Arabidopsis thaliana]
NCBI Description
                   4816
Seq. No.
Contig ID
                   4074 3.R1040
5'-most EST
                  LIB3170-012-Q1-J1-F9
                   4817
Seq. No.
                   4075 1.R1040
Contig ID
5'-most EST
                   zhf700951788.h1
Method
                   BLASTX
                   g1293835
NCBI GI
BLAST score
                   486
                   2.0e-48
E value
                   286
Match length
% identity
                  (U56965) C15H9.5 gene product [Caenorhabditis elegans]
NCBI Description
Seq. No.
                   4818
                   4075 3.R1040
Contig ID
5'-most EST
                   LIB3170-021-Q1-K1-B9
                   4819
Seq. No.
                   4075 4.R1040
Contig ID
                   jC-gmst02400071d07d1
5'-most EST
                   BLASTN
Method
                   g3510343
NCBI GI
BLAST score
                   45
                   6.0e-16
E value
Match length
                   141
% identity
                   83
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MJC20, complete sequence [Arabidopsis thaliana]
                   4820
Seq. No.
                   4076 1.R1040
Contig ID
                   LIB3039-028-Q1-E1-A2
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4127456
BLAST score
                   321
```

2.0e-29 E value Match length 149 % identity

(AJ010818) Cpn21 protein [Arabidopsis thaliana] NCBI Description

Seq. No. 4821



```
4076 2.R1040
Contig ID
                  zhf700958459.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4127456
BLAST score
                  222
                  5.0e-18
E value
Match length
                  87
% identity
                  43
                  (AJ010818) Cpn21 protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4822
                  4077 1.R1040
Contig ID
5'-most EST
                  fC-gmro700750047a2
Method
                  BLASTX
NCBI GI
                  q422337
BLAST score
                  154
E value
                  6.0e-10
Match length
                  55
                  28
% identity
                  protein disulfide-isomerase homolog precursor - fluke
NCBI Description
                  (Schistosoma mansoni) >qi 312018 emb CAA80520 (Z22933)
                  protein disulfide isomerase homologue [Schistosoma mansoni]
Seq. No.
                  4823
Contig ID
                  4077 2.R1040
5'-most EST
                  LIB3106-098-Q1-K1-G1
Method
                  BLASTX
NCBI GI
                  q422337
BLAST score
                  230
                  6.0e-19
E value
Match length
                  67
% identity
                  34
                  protein disulfide-isomerase homolog precursor - fluke
NCBI Description
                  (Schistosoma mansoni) >gi 312018 emb CAA80520 (Z22933)
                  protein disulfide isomerase homologue [Schistosoma mansoni]
Seq. No.
                  4824
                  4077 3.R1040
Contig ID
5'-most EST
                  LIB3065-012-Q1-N1-G7
                  4825
Seq. No.
                  4078 1.R1040
Contig ID
5'-most EST
                  uC-gmropic060a12b1
                  BLASTX
Method
NCBI GI
                  g3522961
BLAST score
                  1192
E value
                  1.0e-154
                  357
Match length
                  79
% identity
NCBI Description
                  (AC004411) putative pto kinase [Arabidopsis thaliana]
Seq. No.
                  4826
                  4078 2.R1040
Contig ID
                  uC-gmrominsoy153c05b1
5'-most EST
```

BLASTN

51

q1216483

Method NCBI GI

BLAST score



```
2.0e-19
E value
                   143
Match length
% identity
                   84
                  Arabidopsis thaliana dual specificity kinase 1 (ADK1) mRNA,
NCBI Description
                   complete cds
                   4827
Seq. No.
                   4078 3.R1040
Contig ID
                   uC-gmropic036f07b1
5'-most EST
Method
                   BLASTX
                   g3522961
NCBI GI
                   379
BLAST score
                   2.0e-36
E value
                   93
Match length
% identity
                   78
                   (AC004411) putative pto kinase [Arabidopsis thaliana]
NCBI Description
                   4828
Seq. No.
                   4078 4.R1040
Contig ID
                   LIB3107-010-Q1-K1-A10
5'-most EST
Method
                   BLASTX
                   q3668069
NCBI GI
                   378
BLAST score
                   3.0e-36
E value
                   109
Match length
% identity
                   65
                   (U28007) Pto kinase interactor 1 [Lycopersicon esculentum]
NCBI Description
                   4829
Seq. No.
                   4078 5.R1040
Contig ID
                   LIB3074-018-Q1-E1-F7
5'-most EST
                   BLASTX
Method
                   g1586940
NCBI GI
                   874
BLAST score
                   5.0e-94
E value
                   248
Match length
                   71
% identity
NCBI Description Ser/Thr kinase [Lycopersicon esculentum]
                   4830
Seq. No.
                   4078 6.R1040
Contig ID
                   LIB3050-025-Q1-K1-D1
5'-most EST
Method
                   BLASTX
                   g3242705
NCBI GI
BLAST score
                   1076
                   1.0e-117
E value
Match length
                   235
                   85
% identity
                   (AC003040) putative nicotinate phosphoribosyltransferase
NCBI Description
                   [Arabidopsis thaliana]
```

Seq. No. 4831

Contig ID 4078 7.R1040

5'-most EST LIB3107-037-Q1-K1-G11

Method BLASTX
NCBI GI g3668069
BLAST score 352



E value 4.0e-33 Match length 121 % identity 60

NCBI Description (U28007) Pto kinase interactor 1 [Lycopersicon esculentum]

Seq. No. 4832

Contig ID 4078_8.R1040 5'-most EST txt700731595.h1

Method BLASTX
NCBI GI g3668069
BLAST score 416
E value 9.0e-41
Match length 123
% identity 65

NCBI Description (U28007) Pto kinase interactor 1 [Lycopersicon esculentum]

Seq. No. 4833

Contig ID 4078 9.R1040 5'-most EST V4L-02-Q1-B1-B12

Method BLASTX
NCBI GI g2914700
BLAST score 700
E value 8.0e-74
Match length 169
% identity 84

NCBI Description (AC003974) tRNA-processing protein SEN3-like [Arabidopsis

thaliana]

Seq. No. 4834

Contig ID 4078_10.R1040

5'-most EST LIB3051-077-Q1-K1-B2

Method BLASTX
NCBI GI g2914700
BLAST score 918
E value 2.0e-99
Match length 259
% identity 73

NCBI Description (AC003974) tRNA-processing protein SEN3-like [Arabidopsis

thaliana]

Seq. No. 4835

Contig ID 4078_11.R1040 5'-most EST uC-gmropic068h01b1

Method BLASTX
NCBI GI g3242705
BLAST score 283
E value 5.0e-25
Match length 67
% identity 78

NCBI Description (AC003040) putative nicotinate phosphoribosyltransferase

[Arabidopsis thaliana]

Seq. No. 4836

Contig ID 4078 12.R1040

5'-most EST LIB3\overline{1}39-099-P1-N1-H8

Method BLASTX NCBI GI g1586940



```
BLAST score
E value
                  1.0e-08
                  63
Match length
                  51
% identity
NCBI Description Ser/Thr kinase [Lycopersicon esculentum]
                  4837
Seq. No.
                  4078 13.R1040
Contig ID
5'-most EST
                  LIB3051-049-Q1-K1-D6
                  4838
Seq. No.
                  4078 15.R1040
Contig ID
5'-most EST
                  eep700867417.h1
```

Seq. No. 4839

Contig ID 4080_1.R1040

5'-most EST LIB3028-021-Q1-B1-F5

Seq. No. 4840

Contig ID 4081 1.R1040

5'-most EST LIB3138-083-P1-N1-F4

Seq. No. 4841

Contig ID 4081_2.R1040

5'-most EST LIB3109-053-Q1-K1-A9

Method BLASTX
NCBI GI g118561
BLAST score 217
E value 7.0e-17
Match length 204
% identity 31

NCBI Description GLUCOSE 1-DEHYDROGENASE B

Seq. No. 4842

Contig ID 4081_3.R1040 5'-most EST zzp700835451.h1

Seq. No. 4843

Contig ID 4082 1.R1040 5'-most EST rca701000512.h1

Seq. No. 4844

Contig ID 4082 2.R1040

5'-most EST uC-gmrominsoy236h07b1

Seq. No. 4845

Contig ID 4083_1.R1040 5'-most EST txt700734689.h1

Method BLASTX
NCBI GI g730526
BLAST score 909
E value 3.0e-98
Match length 193
% identity 88

NCBI Description 60S RIBOSOMAL PROTEIN L13 (BBC1 PROTEIN HOMOLOG)

>gi_480787_pir__S37271 ribosomal protein L13 - Arabidopsis
thaliana >gi_404166_emb_CAA53005_ (X75162) BBC1 protein



[Arabidopsis thaliana]

5 MOSC B51 B1B5051 051 Q1

Seq. No. 4847

Contig ID 4089_1.R1040

5'-most EST LIB3028-021-Q1-B1-G4

Seq. No. 4848

Contig ID 4091 1.R1040

5'-most EST LIB3028-021-Q1-B1-G6

Method BLASTX
NCBI GI g4490316
BLAST score 186
E value 6.0e-14
Match length 47
% identity 68

NCBI Description (AL035678) nucellin-like protein [Arabidopsis thaliana]

Seq. No. 4849

Contig ID 4093 1.R1040

5'-most EST LIB3028-021-Q1-B1-G8

Seq. No. 4850

Contig ID 4096 1.R1040

5'-most EST LIB3039-034-Q1-E1-F11

Method BLASTX
NCBI GI g1685005
BLAST score 1014
E value 1.0e-110
Match length 474
% identity 44

NCBI Description (U32644) immediate-early salicylate-induced

glucosyltransferase [Nicotiana tabacum]

Seq. No. 4851

Contig ID 4099_1.R1040 5'-most EST rca701001277.h1

Method BLASTX
NCBI GI g267131
BLAST score 531
E value 1.0e-53
Match length 289
% identity 25

NCBI Description NUCLEOLYSIN TIAR (TIA-1 RELATED PROTEIN)

>gi_423120_pir_ A46174 RNA-binding protein TIAR - human
>gi_189310 (M96954) nucleolysin TIAR [Homo sapiens]
>gi_4507499_ref_NP_003243.1_pTIAL1_ TIA1 cytotoxic

granule-associated RNA-binding protein-like

Seq. No. 4852

Contig ID 4100_1.R1040

5'-most EST uC-gmropic044h04b1

Method BLASTX NCBI GI g4508079



```
279
BLAST score
E value
                   9.0e-25
Match length
                  107
% identity
                  57
                  (AC005882) 66284 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4853
                   4100 2.R1040
Contig ID
5'-most EST
                   vzy700754417.h1
Method
                  BLASTX
                  q4508079
NCBI GI
BLAST score
                  258
                   2.0e-22
E value
Match length
                   96
% identity
                   61
                   (AC005882) 66284 [Arabidopsis thaliana]
NCBI Description
                   4854
Seq. No.
                   4101 1.R1040
Contig ID
                  LIB3040-054-Q1-E1-G11
5'-most EST
Method
                  BLASTX
                   g541968
NCBI GI
BLAST score
                   267
                   6.0e-23
E value
Match length
                   91
                   55
% identity
                  HMG 1 protein - garden pea >gi_436424_emb_CAA54168_
NCBI Description
                   (X76774) HMG 1 protein [Pisum sativum]
                   4855
Seq. No.
                   4103 1.R1040
Contig ID
                   pmv700891270.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3128168
BLAST score
                   450
E value
                   2.0e-44
Match length
                   142
% identity
                   56
                   (AC004521) putative carboxyl-terminal peptidase
NCBI Description
                   [Arabidopsis thaliana]
                   4856
Seq. No.
                   4103 2.R1040
Contig ID
                   LIB3109-034-Q1-K1-F8
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3128168
BLAST score
                   318
                   4.0e-29
E value
Match length
                   105
                   52
% identity
NCBI Description
                   (AC004521) putative carboxyl-terminal peptidase
```

[Arabidopsis thaliana]

4857 Seq. No.

Contig ID 4104 1.R1040

LIB3106-076-Q1-K1-G1 5'-most EST

Method BLASTX



```
NCBI GI g1172874
BLAST score 753
E value 8.0e-80
Match length 289
% identity 58
```

NCBI Description DEHYDRATION-RESPONSIVE PROTEIN RD22 PRECURSOR

>gi_479589_pir__S34823 dehydration-induced protein RD22 Arabidopsis thaliana >gi_391608_dbj_BAA01546_ (D10703) rd22
[Arabidopsis thaliana] >gi_447134_prf__1913421A rd22 gene

[Arabidopsis thaliana]

Seq. No. 4858

Contig ID 4104 2.R1040

5'-most EST LIB3039-044-Q1-E1-A11

Seq. No. 4859

Contig ID 4108 1.R1040

5'-most EST uC-gmflminsoy053c11b1

Method BLASTX
NCBI GI 94325371
BLAST score 914
E value 1.0e-98
Match length 321
% identity 59

NCBI Description (AF128396) contains similarity to Medicago truncatula N7

protein (GB:Y17613) [Arabidopsis thaliana]

Seq. No. 4860

Contig ID 4108 2.R1040

5'-most EST LIB3139-007-P1-N1-F7

Method BLASTX
NCBI GI g4325371
BLAST score 243
E value 2.0e-20
Match length 56
% identity 77

NCBI Description (AF128396) contains similarity to Medicago truncatula N7

protein (GB:Y17613) [Arabidopsis thaliana]

Seq. No. 4861

Contig ID 4108_3.R1040 5'-most EST zzp700836162.h1

Method BLASTX
NCBI GI g4325371
BLAST score 243
E value 2.0e-20
Match length 56
% identity 77

NCBI Description (AF128396) contains similarity to Medicago truncatula N7

protein (GB:Y17613) [Arabidopsis thaliana]

Seq. No. 4862

Contig ID 4108_7.R1040 5'-most EST pxt700945268.h1

Seq. No. 4863

Contig ID 4110 1.R1040



```
gsv701050106.hl
5'-most EST
                  BLASTX
Method
                   g2342719
NCBI GI
BLAST score
                   355
E value
                   2.0e-33
                   184
Match length
% identity
                   46
                  (AC002341) SF16 protein isolog [Arabidopsis thaliana]
NCBI Description
                   4864
Seq. No.
                   4110 2.R1040
Contig ID
                   k117\overline{0}1211941.h1
5'-most EST
                   BLASTX
Method
                   g2342719
NCBI GI
BLAST score
                   315
E value
                   1.0e-28
Match length
                   93
                   66
% identity
                   (AC002341) SF16 protein isolog [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4865
                   4110 3.R1040
Contig ID
                   LIB3093-051-Q1-K1-C12
5'-most EST
                   4866
Seq. No.
                   4110 4.R1040
Contig ID
                   LIB3139-009-P1-N1-A8
5'-most EST
                   4867
Seq. No.
                   4112 1.R1040
Contig ID
                   k11701208153.h1
5'-most EST
                   BLASTX
Method
                   g3024022
NCBI GI
                   779
BLAST score
                   6.0e-83
E value
                   159
Match length
% identity
                   91
                   INITIATION FACTOR 5A-5 (EIF-5A) (EIF-4D)
NCBI Description
                   >gi_2225885_dbj_BAA20879_ (AB004826) eukaryotic initiation
                   factor 5A5 [Solanum tuberosum]
                   4868
Seq. No.
                   4114 2.R1040
Contig ID
                   leu701150411.hl
5'-most EST
                   4869
Seq. No.
                   4116 1.R1040
Contig ID
                   uC-gmropic064d12b1
5'-most EST
Method
                   BLASTX
                   g3287680
NCBI GI
BLAST score
                   260
```

3.0e-22 E value 139 Match length

50 % identity (AC003979) T22J18.4 [Arabidopsis thaliana] NCBI Description

Seq. No. 4870



```
Contig ID
                  4117 1.R1040
                  LIB3028-021-Q1-B1-C12
5'-most EST
                  BLASTX
Method
                  g1931655
NCBI GI
                  360
BLAST score
                  4.0e-34
E value
                  95
Match length
                  74
% identity
                  (U95973) receptor-kinase isolog [Arabidopsis thaliana]
NCBI Description
                  4871
Seq. No.
                  4118 1.R1040
Contig ID
5'-most EST
                  LIB3028-021-Q1-B1-C2
Seq. No.
                   4872
                  4119 1.R1040
Contig ID
5'-most EST
                  epx701110372.h1
                  BLASTX
Method
                  g3033375
NCBI GI
BLAST score
                  268
                  2.0e-23
E value
Match length
                   92
                  57
% identity
                  (AC004238) putative berberine bridge enzyme [Arabidopsis
NCBI Description
                  thaliana]
                   4873
Seq. No.
                   4121 1.R1040
Contig ID
                  leu701153802.h1
5'-most EST
                  BLASTX
Method
                  q4103324
NCBI GI
BLAST score
                   1222
                   1.0e-135
E value
Match length
                   275
% identity
                   87
                  (AF022716) GDP-mannose pyrophosphorylase [Solanum
NCBI Description
                  tuberosum]
                   4874
Seq. No.
Contig ID
                   4121 2.R1040
5'-most EST
                   leu701146422.h1
Method
                   BLASTX
NCBI GI
                   q4103324
BLAST score
                   506
E value
                   2.0e-51
Match length
                   105
% identity
NCBI Description
                   (AF022716) GDP-mannose pyrophosphorylase [Solanum
                   tuberosum]
                   4875
Seq. No.
                   4121 4.R1040
Contig ID
                   zpv700761104.h1
5'-most EST
                   BLASTN
Method
```

940

g4151924

9.0e-39

83

NCBI GI

E value

BLAST score



Match length 239 % identity 84

NCBI Description Arabidopsis thaliana CYT1 protein (cyt1) mRNA, complete cds

Seq. No. 4876

Contig ID 4121_5.R1040 5'-most EST vzy700751791.h1

Method BLASTX
NCBI GI g4103324
BLAST score 406
E value 1.0e-39
Match length 86
% identity 94

NCBI Description (AF022716) GDP-mannose pyrophosphorylase [Solanum

tuberosum]

Seq. No. 4877

Contig ID 4121 6.R1040

5'-most EST uC-qmrominsoy086a10b1

Method BLASTX
NCBI GI g2642159
BLAST score 261
E value 1.0e-22
Match length 55
% identity 95

NCBI Description (AC003000) putative mannose-1-phosphate guanyltransferase

[Arabidopsis thaliana] >gi_3598958 (AF076484) GDP-mannose

pyrophosphorylase [Arabidopsis thaliana] >gi_4151925

(AF108660) CYT1 protein [Arabidopsis thaliana]

Seq. No. 4878

Contig ID 4121_7.R1040 5'-most EST zhf700959268.h1

Method BLASTN
NCBI GI g2642152
BLAST score 41
E value 8.0e-14
Match length 101

% identity 85
NCBI Description Arabidopsis thaliana chromosome II BAC T517 genomic

sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 4879

Contig ID 4123_1.R1040 5'-most EST awf700839701.h1

Method BLASTX
NCBI GI g2129758
BLAST score 641
E value 8.0e-67
Match length 151
% identity 78

NCBI Description ubiquitin conjugating enzyme E2 protein - Arabidopsis

thaliana >gi 992704 (U33757) UBC7 [Arabidopsis thaliana]

Seq. No. 4880

Contig ID 4123 2.R1040

5'-most EST LIB3049-041-Q1-E1-G12

Match length

% identity

116

86



```
BLASTX
Method
                   g2129758
NCBI GI
BLAST score
                   174
E value
                   3.0e-12
Match length
                   34
% identity
                   94
                   ubiquitin conjugating enzyme E2 protein - Arabidopsis
NCBI Description
                   thaliana >qi 992704 (U33757) UBC7 [Arabidopsis thaliana]
Seq. No.
                   4881
Contiq ID
                   4123 3.R1040
5'-most EST
                   gsv701056366.hl
Method
                   BLASTX
NCBI GI
                   g2129758
BLAST score
                   659
E value
                   5.0e-69
Match length
                   152
% identity
                   80
                   ubiquitin conjugating enzyme E2 protein - Arabidopsis
NCBI Description
                   thaliana >gi 992704 (U33757) UBC7 [Arabidopsis thaliana]
Seq. No.
                   4882
Contig ID
                   4123 4.R1040
                   LIB3139-098-P1-N1-C5
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2129758
BLAST score
                   383
                   5.0e-37
E value
Match length
                   88
                   78
% identity
                   ubiquitin conjugating enzyme E2 protein - Arabidopsis
NCBI Description
                   thaliana >gi 992704 (U33757) UBC7 [Arabidopsis thaliana]
                   4883
Seq. No.
                   4123 6.R1040
Contig ID
                   leu701157761.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2624417
BLAST score
                   174
                   1.0e-12
E value
Match length
                   47
                   66
% identity
                   (AJ002959) Ubiquitin carrier protein [Zea mays]
NCBI Description
                   4884
Seq. No.
                   4123 8.R1040
Contig ID
                   uxk7\overline{0}0667391.h1
5'-most EST
                   4885
Seq. No.
                   4124 1.R1040
Contig ID
5'-most EST
                   ncj700981844.h1
Method
                   BLASTX
NCBI GI
                   g1040877
BLAST score
                   572
E value
                   6.0e-59
```



```
NCBI Description
                  (U30460) expansin S2 precursor [Cucumis sativus]
                  4886
Seq. No.
                  4124 2.R1040
Contig ID
5'-most EST
                  jC-gmro02910046c02a1
Method
                  BLASTX
NCBI GI
                  g1040877
BLAST score
                  464
                  2.0e-46
E value
Match length
                  115
% identity
                  72
                  (U30460) expansin S2 precursor [Cucumis sativus]
NCBI Description
                  4887
Seq. No.
                  4125 1.R1040
Contig ID
5'-most EST
                  LIB3028-021-Q1-B1-D12
Method
                  BLASTX
NCBI GI
                  g4218991
BLAST score
                  217
E value
                  2.0e-17
                  82
Match length
% identity
                  51
                  (AF098632) subtilisin-like protease [Arabidopsis thaliana]
NCBI Description
                  4888
Seq. No.
Contig ID
                  4125 2.R1040
5'-most EST
                  pmv700889380.h1
Method
                  BLASTX
NCBI GI
                  q4218991
BLAST score
                  312
E value
                  2.0e-28
Match length
                  129
% identity
                  49
                  (AF098632) subtilisin-like protease [Arabidopsis thaliana]
NCBI Description
                  4889
Seq. No.
Contig ID
                  4126 1.R1040
5'-most EST
                  LIB3040-003-Q1-E1-E4
                  BLASTN
Method
NCBI GI
                  g407800
BLAST score
                  60
E value
                  6.0e-25
Match length
                  80
% identity
                  94
NCBI Description
                  G.hirsutum mRNA for ribosomal protein 41, large subunit
                   (RL41)
                  4890
Seq. No.
Contig ID
                  4126 2.R1040
5'-most EST
                  LIB3170-013-Q1-J1-F4
Method
                  BLASTN
NCBI GI
                  q407800
BLAST score
                  60
                  5.0e-25
E value
Match length
                  80
% identity
                  94
NCBI Description G.hirsutum mRNA for ribosomal protein 41, large subunit
```



(RL41)

4891 Seq. No. Contig ID 4130 1.R1040 5'-most EST LIB3028-020-Q1-B1-H8 Method BLASTX g2098709 NCBI GI BLAST score 716 7.0e-76 E value 182 Match length % identity 71 (U82975) pectinesterase [Citrus sinensis] NCBI Description Seq. No. 4892 Contig ID 4133 1.R1040 5'-most EST LIB3028-021-Q1-B1-A5 Method BLASTX NCBI GI g3242079 BLAST score 223 7.0e-18 E value Match length 67 66 % identity (AJ006984) proline-rich protein [Capsicum annuum] NCBI Description Seq. No. 4893 Contig ID 4133 2.R1040 5'-most EST $ncj7\overline{0}0981439.h1$ Method BLASTX NCBI GI g3242079 BLAST score 226 E value 1.0e-18 Match length 68 % identity 65 NCBI Description (AJ006984) proline-rich protein [Capsicum annuum] 4894 Seq. No. Contig ID 4135 1.R1040 hyd700730983.h1 5'-most EST BLASTX Method NCBI GI g1651904 BLAST score 172 E value 5.0e-12 Match length 93 40 % identity NCBI Description (D90901) hypothetical protein [Synechocystis sp.] Seq. No. 4895 Contig ID 4137 1.R1040 5'-most EST zhf700964788.h1 Seq. No. 4896 4141 1.R1040 Contig ID 5'-most EST leu701154311.h1

Method BLASTN NCBI GI g396818 BLAST score 343 E value 0.0e+00



```
Match length
                   935
% identity
                  86
NCBI Description P.sativum pspor1 mRNA encoding porin
Seq. No.
                   4897
                  4142 1.R1040
Contig ID
5'-most EST
                  ncj700981734.hl
Method
                  BLASTX
NCBI GI
                  g4006854
BLAST score
                  235
E value
                  2.0e-19
Match length
                  102
% identity
                   49
NCBI Description
                  (Z99707) putative protein [Arabidopsis thaliana]
                  4898
Seq. No.
                  4142 2.R1040
Contig ID
                  leu701148704.h1
5'-most EST
Seq. No.
                  4899
                  4147 1.R1040
Contig ID
5'-most EST
                  zhf700953783.h1
Method
                  BLASTX
NCBI GI
                  q3287691
BLAST score
                   483
E value
                   4.0e-48
Match length
                  252
% identity
                   45
                  (AC003979) Contains similarity to RING zinc finger protein
NCBI Description
                  gb X95455 from Gallus gallus. [Arabidopsis thaliana]
                   4900
Seq. No.
                   4147 2.R1040
Contig ID
5'-most EST
                   jC-gmle01810064h11a1
Method
                  BLASTX
                  g3642014
NCBI GI
BLAST score
                  173
                  1.0e-11
E value
Match length
                   49
% identity
                   61
                   (AL031620) similar to Zinc finger, C3HC4 type (RING finger)
NCBI Description
                   [Caenorhabditis elegans]
Seq. No.
                   4901
Contig ID
                   4147 3.R1040
5'-most EST
                  LIB3138-012-Q1-N1-C12
Seq. No.
                   4902
                   4148 1.R1040
Contig ID
5'-most EST
                   asn701138390.h1
```

Method BLASTX
NCBI GI g3377517
BLAST score 219
E value 2.0e-17
Match length 99
% identity 41

NCBI Description (AF073361) nitrate transporter NTL1 [Arabidopsis thaliana]



```
Seq. No.
                   4903
Contig ID
                   4149 1.R1040
5'-most EST
                   leu701146016.hl
Method
                   BLASTX
                   q3184289
NCBI GI
BLAST score
                   334
E value
                   6.0e-31
Match length
                   160
% identity
                   51
                  (AC004136) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   4904
Seq. No.
Contig ID
                   4152 1.R1040
5'-most EST
                   txt700732376.h1
                   4905
Seq. No.
Contig ID
                   4153 1.R1040
5'-most EST
                   jC-gmle01810022h01a1
                   4906
Seq. No.
Contig ID
                   4154 1.R1040
5'-most EST
                   LIB3139-093-P1-N1-E3
Method
                   BLASTX
NCBI GI
                   g2982303
BLAST score
                   568
E value
                   2.0e-58
                  204
Match length
% identity
                   58
NCBI Description
                  (AF051236) hypothetical protein [Picea mariana]
                   4907
Seq. No.
                   4154 2.R1040
Contig ID
5'-most EST
                   dpv701103307.h1
Method
                   BLASTX
NCBI GI
                   q2982303
BLAST score
                   539
                   3.0e-55
E value
                   134
Match length
% identity
                   79
                  (AF051236) hypothetical protein [Picea mariana]
NCBI Description
                   4908
Seq. No.
Contig ID
                   4160 1.R1040
5'-most EST
                   LIB3087-005-Q1-K1-G5
Method
                   BLASTX
NCBI GI
                   g4454026
BLAST score
                   903
                   2.0e-97
E value
Match length
                   323
% identity
                   (AL035394) phosphatase like protein [Arabidopsis thaliana]
NCBI Description
```

Seq. No. 4909

Contig ID 4160_2.R1040

5'-most EST uC-gmrominsoy025a07b1

Method BLASTX



```
NCBI GI
                  g4454026
BLAST score
                  379
E value
                  2.0e-36
Match length
                  101
% identity
                  76
                  (AL035394) phosphatase like protein [Arabidopsis thaliana]
NCBI Description
                  4910
Seq. No.
                  4163 1.R1040
Contig ID
5'-most EST
                  sat701010905.h1
Method
                  BLASTX
                  q3297891
NCBI GI
BLAST score
                  934
                  1.0e-101
E value
Match length
                  194
% identity
                  89
                  (AJ002990) nucleotide repair protein [Lilium longiflorum]
NCBI Description
Seq. No.
                  4911
Contig ID
                  4166 1.R1040
5'-most EST
                  uC-gmrominsoy271c10b1
Method
                  BLASTX
NCBI GI
                  g1946355
BLAST score
                  191
E value
                  2.0e-14
Match length
                  119
% identity
                  39
NCBI Description
                  (U93215) maize transposon MuDR mudrA protein isolog
                   [Arabidopsis thaliana] >gi 2880040 (AC002340) maize
                  transposon MuDR mudrA-like protein [Arabidopsis thaliana]
                  4912
Seq. No.
Contig ID
                  4167 1.R1040
5'-most EST
                  zpv700761920.hl
Method
                  BLASTX
NCBI GI
                  q3763927
BLAST score
                  894
                  3.0e-96
E value
Match length
                  287
% identity
                  62
NCBI Description
                   (AC004450) putative carboxyphosphoenolpyruvate mutase
                   [Arabidopsis thaliana]
Seq. No.
                  4913
Contig ID
                  4167 2.R1040
5'-most EST
                  gsv701049034.h1
Method
                  BLASTX
NCBI GI
                  q3763927
BLAST score
                  353
E value
                  2.0e-33
                  108
```

Match length % identity 69

(AC004450) putative carboxyphosphoenolpyruvate mutase NCBI Description

[Arabidopsis thaliana]

Seq. No. 4914

4169_1.R1040 Contig ID



```
5'-most EST
                  LIB3040-033-Q1-E1-C8
Seq. No.
                  4915
Contig ID
                  4171 1.R1040
5'-most EST
                  LIB3138-031-Q1-N1-G6
                  4916
Seq. No.
                  4172 1.R1040
Contig ID
5'-most EST
                  LIB3039-005-Q1-E1-H8
                  4917
Seq. No.
Contig ID
                  4177 1.R1040
5'-most EST
                  dpv701103646.h1
Seq. No.
                  4918
Contig ID
                  4178 1.R1040
5'-most EST
                  jex700909168.h1
                  4919
Seq. No.
                  4180 1.R1040
Contig ID
5'-most EST
                  rlr700900441.h1
                  4920
Seq. No.
                  4182 1.R1040
Contig ID
5'-most EST
                  LIB3109-006-Q1-K1-A6
Method
                  BLASTX
NCBI GI
                  g3328240
BLAST score
                  597
                  1.0e-61
E value
Match length
                  255
% identity
                  47
                  (AF064775) early nodule-specific protein [Medicago
NCBI Description
                  truncatula]
                  4921
Seq. No.
Contig ID
                  4182 2.R1040
5'-most EST
                  LIB3094-055-Q1-K1-E9
Method
                  BLASTX
NCBI GI
                  q1110502
BLAST score
                  307
E value
                  5.0e-28
Match length
                  118
% identity
                  53
NCBI Description
                  (U40387) coil protein [Medicago sativa]
                  4922
Seq. No.
Contig ID
                  4184 1.R1040
                  rca701002018.h1
5'-most EST
Method
                  BLASTX
                  g2398521
NCBI GI
BLAST score
                   526
E value
                  2.0e-53
Match length
                  209
                  53
% identity
```

Seq. No. 4923

NCBI Description

(Y13720) transcription factor [Arabidopsis thaliana]



```
4184 2.R1040
Contig ID
                   ssr700555336.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2398521
BLAST score
                   439
                   5.0e-43
E value
                   223
Match length
% identity
                   48
                  (Y13720) transcription factor [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4924
                   4186 1.R1040
Contig ID
                   fde700873486.h1
5'-most EST
Seq. No.
                   4925
Contig ID
                   4187 1.R1040
5'-most EST
                   taw700659693.h1
Method
                   BLASTX
                   g2695711
NCBI GI
BLAST score
                   360
                   7.0e-34
E value
Match length
                   93
% identity
                   67
                   (AJ001370) cytochome b5 [Olea europaea]
NCBI Description
Seq. No.
                   4926
                   4187 2.R1040
Contig ID
5'-most EST
                   jsh7\overline{0}1070289.h1
Method
                   BLASTX
NCBI GI
                   q2695711
BLAST score
                   372
                   1.0e-35
E value
                   113
Match length
% identity
                   59
                  (AJ001370) cytochome b5 [Olea europaea]
NCBI Description
                   4927
Seq. No.
Contig ID
                   4187 3.R1040
5'-most EST
                   jC-gmro02800033d10a1
                   BLASTX
Method
NCBI GI
                   q729252
BLAST score
                   270
E value
                   5.0e-24
Match length
                   64
% identity
                   73
                   CYTOCHROME B5 >gi_167140 (M87514) cytochrome b-5 [Brassica
NCBI Description
                   oleracea] >gi_384338_prf__1905426A cytochrome b5 [Brassica
                   oleracea]
                   4928
Seq. No.
Contig ID
                   4189 1.R1040
5'-most EST
                   fua701042753.h1
```

Method BLASTX
NCBI GI g4006932
BLAST score 180
E value 5.0e-13
Match length 135



```
% identity
NCBI Description
                  (AJ011400) NADH: ubiquinone oxidoreductase b17.2 subunit
                   [Bos taurus]
                  4929
Seq. No.
Contig ID
                  4189 3.R1040
5'-most EST
                  jC-gmf102220064b02a1
Method
                  BLASTX
NCBI GI
                  g4580395
BLAST score
                  283
                  5.0e-25
E value
Match length
                  88
% identity
                  64
NCBI Description
                  (AC007171) putative kinesin-related protein [Arabidopsis
                  thaliana]
                  4930
Seq. No.
Contig ID
                  4190 1.R1040
5'-most EST
                  vzy700753930.h1
Method
                  BLASTX
NCBI GI
                  g4049354
BLAST score
                  1999
                  0.0e+00
E value
Match length
                  452
% identity
                  83
NCBI Description
                  (AL034567) glycine hydroxymethyltransferase (EC
                  2.1.2.1) - like protein [Arabidopsis thaliana]
                  4931
Seq. No.
Contig ID
                  4190 2.R1040
5'-most EST
                  LIB3107-056-Q1-K1-H10
Seq. No.
                  4932
Contig ID
                  4194_1.R1040
5'-most EST
                  uC-gmflminsoy049d03b1
Method
                  BLASTX
NCBI GI
                  q282994
BLAST score
                  1032
E value
                  1.0e-113
Match length
                  226
% identity
                  81
                  Sip1 protein - barley >gi 167100 (M77475) seed imbibition
NCBI Description
                  protein [Hordeum vulgare]
Seq. No.
                  4933
Contig ID
                  4196 1.R1040
5'-most EST
                  LIB3039-047-Q1-E1-A3
Method
                  BLASTX
NCBI GI
                  q4468813
BLAST score
                  736
```

7.0e-78 E value 247 Match length % identity

NCBI Description

(AL035601) putative protein [Arabidopsis thaliana]

Seq. No.

4934

Contig ID

4196 2.R1040

Method

NCBI GI

BLASTX

g1518388



```
5'-most EST
                  fC-gmse7000751282d1
Seq. No.
                  4935
Contig ID
                  4196 3.R1040
5'-most EST
                  LIB3039-028-Q1-E1-C4
Method
                  BLASTX
                  g4468813
NCBI GI
BLAST score
                  369
E value
                  4.0e-35
Match length
                  119
% identity
                  55
                 (AL035601) putative protein [Arabidopsis thaliana]
NCBI Description
                  4936
Seq. No.
Contig ID
                  4201 1.R1040
5'-most EST
                  jC-qmst02400052e06a1
Seq. No.
                  4937
Contig ID
                  4203 1.R1040
                  uC-gmflminsoy091g11b1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q3461821
BLAST score
                  522
E value
                  2.0e-66
Match length
                  267
% identity
                  54
                  (AC004138) putative nucleoside triphosphatase [Arabidopsis
NCBI Description
                  thaliana]
                   4938
Seq. No.
                   4203 3.R1040
Contig ID
5'-most EST
                  xpa700793151.h1
Method
                  BLASTX
                  g3461821
NCBI GI
BLAST score
                  576
                  2.0e-59
E value
Match length
                  173
% identity
                   62
                  (AC004138) putative nucleoside triphosphatase [Arabidopsis
NCBI Description
                  thaliana]
                   4939
Seq. No.
                   4204 1.R1040
Contig ID
5'-most EST
                   zhf700952486.h1
Method
                   BLASTX
                   q1854581
NCBI GI
                   886
BLAST score
                   2.0e-95
E value
Match length
                   276
% identity
                   64
NCBI Description (L24120) peroxidase precursor [Linum usitatissimum]
                   4940
Seq. No.
Contig ID
                   4204 2.R1040
5'-most EST
                   LIB3106-022-Q1-K1-F10
```

951



4941

BLAST score 163 E value 3.0e-11 Match length 65 % identity 51

NCBI Description (X91172) korean-radish isoperoxidase [Raphanus sativus]

Seq. No.

Contig ID 4205_1.R1040 5'-most EST leu701146640.h1

Method BLASTX
NCBI GI g2129605
BLAST score 932
E value 1.0e-104
Match length 210
% identity 87

NCBI Description GTP-binding protein 2 - Arabidopsis thaliana

>gi_2129702_pir__S71585 Rab2 homolog GTP-binding protein
ATGB2 - Arabidopsis thaliana >gi_1184983 (U46925) ATGB2
[Arabidopsis thaliana] >gi_3805852_emb_CAA21472_ (AL031986)

GTP-binding protein GB2 [Arabidopsis thaliana]

Seq. No. 4942

Contig ID 4207_1.R1040 5'-most EST leu701149522.h1

Method BLASTN
NCBI GI g603218
BLAST score 504
E value 0.0e+00
Match length 884
% identity 89

NCBI Description Medicago sativa glucose-6-phosphate dehydrogenase mRNA,

complete cds

Seq. No. 4943

Contig ID 4207_2.R1040 5'-most EST kl1701215133.h1

Method BLASTX
NCBI GI g3617770
BLAST score 752
E value 1.0e-79
Match length 183
% identity 80

NCBI Description (Y14329) threonyl-tRNA synthetase [Arabidopsis thaliana]

Seq. No. 4944

Contig ID 4207_3.R1040

5'-most EST jC-gmro02910062c06a1

Method BLASTN
NCBI GI g603218
BLAST score 113
E value 1.0e-56
Match length 261
% identity 86

NCBI Description Medicago sativa glucose-6-phosphate dehydrogenase mRNA,

complete cds

Seq. No. 4945

Seq. No.

4950



```
4207 4.R1040
Contig ID
5'-most EST
                   zsg701128628.h1
Method
                   BLASTX
NCBI GI
                   g3617770
BLAST score
                   2232
                   0.0e+00
E value
                   505
Match length
% identity
                   80
                  (Y14329) threonyl-tRNA synthetase [Arabidopsis thaliana]
NCBI Description
                   4946
Seq. No.
                   4207 5.R1040
Contig ID
                   LIB3170-061-Q1-K2-F9
5'-most EST
Method
                   BLASTN
NCBI GI
                   g603218
BLAST score
                   56
                   1.0e-22
E value
Match length
                   116
% identity
                   88
                  Medicago sativa glucose-6-phosphate dehydrogenase mRNA,
NCBI Description
                   complete cds
                   4947
Seq. No.
                   4207 7.R1040
Contig ID
5'-most EST
                   leu7\overline{0}1157358.h1
Method
                   BLASTN
                   g603218
NCBI GI
BLAST score
                   53
                   5.0e-21
E value
Match length
                   77
% identity
                   92
NCBI Description
                  Medicago sativa glucose-6-phosphate dehydrogenase mRNA,
                   complete cds
Seq. No.
                   4948
Contig ID
                   4208 1.R1040
5'-most EST
                   hyd700730986.hl
Method
                   BLASTX
NCBI GI
                   q4091806
BLAST score
                   331
E value
                   6.0e-34
Match length
                   92
% identity
                   53
NCBI Description
                   (AF052585) CONSTANS-like protein 2 [Malus domestica]
                   4949
Seq. No.
Contig ID
                   4210 1.R1040
                   LIB3049-001-Q1-E1-B5
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3599968
BLAST score
                   1008
E value
                   1.0e-110
Match length
                   250
% identity
NCBI Description
                  (AF032123) clp protease [Arabidopsis thaliana]
```

953



```
Contig ID
                  4212 1.R1040
5'-most EST
                  LIB3050-016-Q1-E1-C3
Method
                  BLASTX
NCBI GI
                  g4572674
BLAST score
                  209
E value
                  2.0e-16
Match length
                  91
% identity
                  56
NCBI Description
                  (AC006954) unknown protein [Arabidopsis thaliana]
                  4951
Seq. No.
                  4212 2.R1040
Contig ID
5'-most EST
                  LIB3050-015-Q1-E1-B8
Seq. No.
                  4952
Contig ID
                  4214 1.R1040
5'-most EST
                  jC-gmst02400031f01a1
Method
                  BLASTX
NCBI GI
                  g2982465
BLAST score
                  200
E value
                  5.0e-15
Match length
                  209
% identity
                  26
NCBI Description
                  (AL022223) putative protein [Arabidopsis thaliana]
                  4953
Seq. No.
Contig ID
                  4215 1.R1040
                  LIB3028-019-Q1-B1-F1
5'-most EST
Method
                  BLASTX
NCBI GI
                  q4558678
BLAST score
                  233
E value
                  2.0e-19
Match length
                  92
                  47
% identity
NCBI Description
                  (AC006586) unknown protein [Arabidopsis thaliana]
                  4954
Seq. No.
Contig ID
                  4215 2.R1040
5'-most EST
                  gsf700698478.hl
Method
                  BLASTX
NCBI GI
                  g4558678
BLAST score
                  141
                  9.0e-09
E value
Match length
                  74
% identity
                  42
NCBI Description
                  (AC006586) unknown protein [Arabidopsis thaliana]
                  4955
Seq. No.
Contig ID
                  4216 1.R1040
5'-most EST
                  LIB3028-019-Q1-B1-F10
Seq. No.
                  4956
```

Contig ID 4223 1.R1040

5'-most EST LIB3028-019-Q1-B1-F8

Seq. No. 4957

Contig ID 4224_1.R1040

BLAST score

Match length

E value

529

154

1.0e-53



```
pxt700945114.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3193306
BLAST score
                   366
E value
                   9.0e-35
Match length
                   104
% identity
NCBI Description
                   (AF069300) contains similarity to Arabidopsis
                   membrane-associated salt-inducible-like protein
                   (GB:AL021637) [Arabidopsis thaliana]
Seq. No.
                   4958
Contig ID
                   4225 1.R1040
5'-most EST
                   LIB3170-035-Q1-K1-D6
Method
                   BLASTN
NCBI GI
                   q1230613
BLAST score
                   116
E value
                   2.0e-58
Match length
                   428
% identity
                   86
NCBI Description
                   Lupinus albus isoflavone reductase-like protein mRNA,
                   complete cds
Seq. No.
                   4959
Contig ID
                   4227 1.R1040
5'-most EST
                   leu701157591.h1
Method
                   BLASTN
NCBI GI
                   q512399
BLAST score
                   399
E value
                   0.0e + 00
Match length
                   923
% identity
                   86
NCBI Description M.sativa mRNA for annexin
                   4960
Seq. No.
Contig ID
                   4227 2.R1040
5'-most EST
                   smc7\overline{0}0747149.h1
Seq. No.
                   4961
Contig ID
                   4228 1.R1040
5'-most EST
                   LIB3109-054-Q1-K1-F1
Method
                   BLASTX
NCBI GI
                   q4455232
BLAST score
                   1938
E value
                   0.0e+00
Match length
                   609
% identity
                   62
NCBI Description
                   (AL035523) putative protein [Arabidopsis thaliana]
Seq. No.
                   4962
Contig ID
                   4229 1.R1040
5'-most EST
                   g5753542
Method
                   BLASTX
NCBI GI
                   q2191138
```

% identity

71



```
% identity
NCBI Description
                  (AF007269) A IG002N01.18 gene product [Arabidopsis
                  thaliana]
Seq. No.
                  4963
Contig ID
                  4229 2.R1040
5'-most EST
                  LIB3106-013-Q1-K1-B12
Method
                  BLASTX
NCBI GI
                  q2191138
BLAST score
                  439
E value
                  2.0e-43
Match length
                  131
% identity
                  (AF007269) A IG002N01.18 gene product [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  4964
Contig ID
                  4229 3.R1040
5'-most EST
                  LIB3106-051-Q1-K1-E1
Seq. No.
                  4965
Contig ID
                  4230 1.R1040
5'-most EST
                  leu701146067.h1
Method
                  BLASTX
NCBI GI
                  g2702281
BLAST score
                  1132
E value
                  1.0e-124
Match length
                  276
% identity
                  77
NCBI Description
                  (AC003033) putative protein disulfide isomerase precursor
                  [Arabidopsis thaliana]
Seq. No.
                  4966
                  4234 1.R1040
Contig ID
5'-most EST
                  fde700877068.hl
Method
                  BLASTX
NCBI GI
                  g1651651
BLAST score
                  241
                  5.0e-41
E value
Match length
                  179
% identity
                  53
NCBI Description (D90899) prenyltransferase [Synechocystis sp.]
                  4967
Seq. No.
                  4234 2.R1040
Contig ID
                  LIB3055-013-Q1-N1-E3
5'-most EST
                  4968
Seq. No.
                  4237 1.R1040
Contig ID
5'-most EST
                  LIB3056-014-Q1-N1-A11
Method
                  BLASTX
                  g3367534
NCBI GI
BLAST score
                  1798
E value
                  0.0e + 00
Match length
                  500
```

NCBI Description (AC004392) Strong similarity to coatamer alpha subunit



(HEPCOP) homolog gb_U24105 from Homo sapiens. [Arabidopsis thaliana]

Seq. No. 4969

Contig ID 4240 1.R1040

5'-most EST LIB3170-082-Q1-K1-D12

Method BLASTX NCBI GI q4490332 BLAST score 380 E value 2.0e-36 Match length 124 % identity 54

(AL035656) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No.

4970

Contig ID 4240 2.R1040

5'-most EST LIB3028-019-Q1-B1-D9

Method BLASTX NCBI GI q4490332 BLAST score 331 E value 3.0e-37 Match length 121 % identity

(AL035656) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 4971

Contig ID 4241 1.R1040 5'-most EST hyd700730638.h1

Method BLASTX NCBI GI g2257531 BLAST score 166 E value 8.0e-11 Match length 165 % identity 28

(AB004537) WEB1 PROTEIN [Schizosaccharomyces pombe] NCBI Description

>gi_2950507_emb_CAA17835 (AL022072) web1 homolog; protein transport protein; WD-repeat protein [Schizosaccharomyces

pombe]

4972 Seq. No.

Contig ID 4241 2.R1040

5'-most EST LIB3049-015-Q1-E1-G6

4973 Seq. No.

4241 3.R1040 Contig ID 5'-most EST leu701155661.h1

Method BLASTX NCBI GI q3845237 BLAST score 181 E value 3.0e-13Match length 102 % identity 39

NCBI Description (AE001407) WEB-1 ortholog, WD40 [Plasmodium falciparum]

Seq. No.

4974

Contig ID 4244 1.R1040

5'-most EST LIB3028-019-Q1-B1-E3



Seq. No. 4975

Contig ID 4244_2.R1040 5'-most EST leu701157087.h1

Seq. No. 4976

Contig ID 4245 1.R1040

5'-most EST LIB3050-011-Q1-E1-B9

Method BLASTN
NCBI GI g287810
BLAST score 162
E value 1.0e-85
Match length 241
% identity 94

NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A

Seq. No. 4977

Contig ID 4245_2.R1040 5'-most EST uxk700671425.h1

Method BLASTN
NCBI GI g287810
BLAST score 316
E value 1.0e-177
Match length 558
% identity 91

NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A

Seq. No. 4978

Contig ID 4245_3.R1040 5'-most EST rlr700900452.h1

Method BLASTN
NCBI GI g287810
BLAST score 449
E value 0.0e+00
Match length 705
% identity 91

NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A

Seq. No. 4979

Contig ID 4245_4.R1040

5'-most EST LIB3087-012-Q1-K1-F4

Method BLASTN
NCBI GI g287810
BLAST score 123
E value 1.0e-62
Match length 221
% identity 90

NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A

Seq. No. 4980

Contig ID 4245 5.R1040

5'-most EST LIB3170-075-Q1-K2-A8

Method BLASTN
NCBI GI g287810
BLAST score 174
E value 5.0e-93
Match length 318

% identity

83



% identity NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A 4981 Seq. No. 4245 6.R1040 Contig ID LIB3109-011-Q1-K1-G1 5'-most EST BLASTN Method g287810 NCBI GI BLAST score 143 1.0e-74 E value Match length 238 91 % identity NCBI Description M.sativa mRNA for phosphoprotein phosphatase type 2A 4982 Seq. No. 4246 1.R1040 Contig ID 5'-most EST LIB3028-019-Q1-B1-C9 Method BLASTX q4191791 NCBI GI BLAST score 356 6.0e - 34E value Match length 82 77 % identity (AC005917) putative sf21 {Helianthus annuus} protein NCBI Description [Arabidopsis thaliana] 4983 Seq. No. 4251_1.R1040 Contig ID 5'-most EST epx701107616.h1Method BLASTX NCBI GI q2052383 BLAST score 1559 1.0e-174 E value Match length 347 % identity 79 NCBI Description (U66345) calreticulin [Arabidopsis thaliana] 4984 Seq. No. 4251 2.R1040 Contig ID 5'-most EST LIB3107-015-Q1-K1-F9 Method BLASTX NCBI GI g2052383 BLAST score 256 E value 4.0e-22 Match length 54 83 % identity NCBI Description (U66345) calreticulin [Arabidopsis thaliana] Seq. No. 4985 Contig ID 4253 1.R1040 jex700905636.h1 5'-most EST Method BLASTN g2970653 NCBI GI BLAST score 199 E value 1.0e-107 Match length 575



NCBI Description Vigna unguiculata ferritin subunit cowpea2 precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds

Seq. No. 4986

Contig ID 4253_3.R1040 5'-most EST fde700875371.h1

Method BLASTN
NCBI GI g2970653
BLAST score 37
E value 2.0e-11
Match length 97

Match length 97 % identity 85

NCBI Description Vigna unguiculata ferritin subunit cowpea2 precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds

Seq. No. 4987

Contig ID 4260 1.R1040

5'-most EST LIB3028-019-Q1-B1-C2

Method BLASTX
NCBI GI g1055130
BLAST score 179
E value 4.0e-13
Match length 85
% identity 44

NCBI Description (U39998) coded for by C. elegans cDNA yk92b11.3; coded for

by C. elegans cDNA yk92b11.5; coded for by C. elegans cDNA yk78c2.5; coded for by C. elegans cDNA cm9a8; coded for by C. elegans cDNA yk66h8.3; coded for by C. elegans cDNA

yk78c2.3

Seq. No. 4988

Contig ID 4269 1.R1040

5'-most EST uC-gmronoir018f07b1

Seq. No. 4989

Contig ID 4269_2.R1040 5'-most EST fde700874154.h1

Seq. No. 4990

Contig ID 4275_1.R1040 5'-most EST sat701003283.h1

Method BLASTX
NCBI GI g2129933
BLAST score 699
E value 9.0e-74
Match length 130
% identity 96

NCBI Description myb-related transcription factor TMH27 - tomato

>gi_1167484_emb_CAA64614 (X95296) transcription factor

[Lycopersicon esculentum]

Seq. No. 4991

Contig ID 4276_1.R1040 5'-most EST crh700854095.h1

Seq. No. 4992

Contig ID 4278 1.R1040



```
5'-most EST
                   fua701042058.h1
Method
                   BLASTX
NCBI GI
                   g3877252
BLAST score
                   306
E value
                   7.0e-28
Match length
                   137
% identity
                   47
                   (Z93382) F45G2.10 [Caenorhabditis elegans]
NCBI Description
Seq. No.
                   4993
Contig ID
                   4282 1.R1040
5'-most EST
                   gsv701045871.h1
Method
                   BLASTX
NCBI GI
                   g3152660
BLAST score
                   361
E value
                   1.0e-33
Match length
                   322
% identity
                   29
                  (AF064603) GA17 protein [Homo sapiens]
NCBI Description
Seq. No.
                   4994
                   4282 2.R1040
Contig ID
5'-most EST
                   pcp700995151.h1
                   4995
Seq. No.
Contig ID
                   4282 4.R1040
5'-most EST
                   zsg7\overline{0}1121632.h1
                   4996
Seq. No.
                   4287 1.R1040
Contig ID
                   eep7\overline{0}0868521.h1
5'-most EST
Seq. No.
                   4997
Contig ID
                   4291 1.R1040
5'-most EST
                   LIB3028-010-Q1-B1-H2
Method
                   BLASTX
NCBI GI
                   g2129858
BLAST score
                   465
E value
                   1.0e-46
Match length
                   110
% identity
                   75
NCBI Description
                   trypsin inhibitor precursor (clone ATI18) - alfalfa
                   >gi 509372 emb CAA56253 (X79879) serine proteinase
                   inhibitor [Medicago sativa]
                   4998
Seq. No.
Contig ID
                   4298 1.R1040
5'-most EST
                   LIB3073-004-Q1-K1-B7
Method
                   BLASTX
NCBI GI
                   q2500345
BLAST score
                   444
E value
                   7.0e-44
```

Match length 128 % identity 70

NCBI Description NHP2/RS6 FAMILY PROTEIN YEL026W HOMOLOG (HIGH MOBILITY

GROUP-LIKE NUCLEAR PROTEIN 2 HOMOLOG)

>gi_2618578_dbj BAA23363 (D50420) OTK27 [Homo sapiens]



>gi_3859990 (AF091076) OTK27 [Homo sapiens]
>gi_1589072_prf__2210268A nuclear protein-NHP2-like protein
[Homo sapiens]

Seq. No. 4999

Contig ID 4299 1.R1040

5'-most EST LIB3030-010-Q1-B1-E4

Method BLASTX
NCBI GI g619747
BLAST score 649
E value 7.0e-68
Match length 182
% identity 65

NCBI Description (U18968) phosphoribosylanthranilate isomerase [Arabidopsis

thaliana]

Seq. No. 5000

Contig ID 4300 1.R1040

5'-most EST LIB3028-018-Q1-B1-E2

Seq. No. 5001

Contig ID 4300 2.R1040

5'-most EST jC-gmf102220098b05d1

Seq. No. 5002

Contig ID 4301 1.R1040

5'-most EST LIB3170-076-Q1-K1-A10

Method BLASTX
NCBI GI g4325282
BLAST score 609
E value 3.0e-63
Match length 146
% identity 74

NCBI Description (AF123310) NAC domain protein NAM [Arabidopsis thaliana]

>gi_4325286_gb_AAD17314_ (AF123311) NAC domain protein NAM

[Arabidopsis thaliana]

Seq. No. 5003

Contig ID 4302 1.R1040 5'-most EST pcp700994416.h1

Method BLASTN
NCBI GI g4519263
BLAST score 267
E value 1.0e-148
Match length 583
% identity 86

NCBI Description Citrus unshiu mRNA for vacuolar H+-ATPase B subunit,

complete cds

Seq. No. 5004

Contig ID 4302_2.R1040 5'-most EST zzp700836191.h1

Method BLASTX
NCBI GI g2493129
BLAST score 2395
E value 0.0e+00
Match length 489



% identity 96

NCBI Description VACUOLAR ATP SYNTHASE SUBUNIT B ISOFORM 1 (V-ATPASE B SUBUNIT) >gi 459198 (U07052) vacuolar H+-ATPase subunit B

[Gossypium hirsutum]

Seq. No. 5005

Contig ID 4302 3.R1040

5'-most EST LIB3040-044-Q1-E1-A7

Method BLASTN
NCBI GI g459197
BLAST score 152
E value 7.0e-80
Match length 348
% identity 86

NCBI Description Gossypium hirsutum vacuolar H+-ATPase subunit B mRNA,

complete cds

Seq. No. 5006

4302 4.R1040 Contig ID 5'-most EST q5752591 Method BLASTN q4519263 NCBI GI BLAST score 92 4.0e-44 E value 244 Match length % identity 84

NCBI Description Citrus unshiu mRNA for vacuolar H+-ATPase B subunit,

complete cds

Seq. No. 5007

Contig ID 4302_5.R1040 5'-most EST hyd70729925.h1

Method BLASTN
NCBI GI 9459197
BLAST score 72
E value 3.0e-32
Match length 184
% identity 85

NCBI Description Gossypium hirsutum vacuolar H+-ATPase subunit B mRNA,

complete cds

Seq. No. 5008

Contig ID 4302_8.R1040 5'-most EST zhf700960468.h1

Method BLASTN
NCBI GI g459197
BLAST score 44
E value 1.0e-15
Match length 60
% identity 93

NCBI Description Gossypium hirsutum vacuolar H+-ATPase subunit B mRNA,

complete cds

Seq. No. 5009

Contig ID 4310_1.R1040 5'-most EST leu701145010.h1

Method BLASTX



```
g4038044
NCBI GI
                  222
BLAST score
E value
                   6.0e-18
Match length
                  83
% identity
NCBI Description
                   (AC005936) unknown protein [Arabidopsis thaliana]
                  >gi 4406788 gb AAD20098 (AC006532) unknown protein
                   [Arabidopsis thaliana]
                   5010
Seq. No.
                   4318 1.R1040
Contig ID
                  LIB3028-018-Q1-B1-C4
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2245125
BLAST score
                  145
E value
                   5.0e-09
Match length
                   56
                   55
% identity
NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]
                   5011
Seq. No.
Contig ID
                   4321 1.R1040
                   gsv701050109.hl
5'-most EST
                   BLASTX
Method
                   g3184282
NCBI GI
BLAST score
                   391
E value
                   7.0e-38
Match length
                   134
% identity
                  (AC004136) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5012
                   4321 2.R1040
Contig ID
                   rrt700646046.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3184282
BLAST score
                   145
                   2.0e-09
E value
Match length
                   50
% identity
                   64
                  (AC004136) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5013
Seq. No.
                   4322 1.R1040
Contig ID
5'-most EST
                   LIB3028-016-Q1-B1-E5
Method
                   BLASTX
NCBI GI
                   g2633727
BLAST score
                   232
E value
                   7.0e-19
Match length
                   124
% identity
                   44
```

(Z99111) ykrT [Bacillus subtilis] NCBI Description

5014 Seq. No.

4325 1.R1040 Contig ID

5'-most EST LIB3109-013-Q1-K1-D9

Method BLASTX



```
NCBI GI
                   g2618705
BLAST score
                   1273
E value
                   1.0e-141
Match length
                   308
% identity
                   (AC002510) putative ABC transporter, 5' partial
NCBI Description
                   [Arabidopsis thaliana]
                   5015
Seq. No.
Contig ID
                   4326 1.R1040
5'-most EST
                  LIB3028-018-Q1-B1-B7
                  BLASTN
Method
NCBI GI
                   g3702734
BLAST score
                   35
                   7.0e-10
E value
Match length
                   147
% identity
                   86
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MNB8, complete sequence [Arabidopsis thaliana]
                   5016
Seq. No.
Contig ID
                   4327 1.R1040
5'-most EST
                   rlr700896704.h1
Method
                   BLASTX
NCBI GI
                   g2688824
BLAST score
                   410
E value
                   9.0e-40
Match length
                   124
% identity
                   69
                  (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                   armeniaca]
                   5017
Seq. No.
                   4327 2.R1040
Contig ID
                   uC-gmrominsoy260b07b1
5'-most EST
Method
                   BLASTX
                   g2688824
NCBI GI
BLAST score
                   168
                   1.0e-11
E value
Match length
                   40
                   78
% identity
                   (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                   armeniaca]
                   5018
Seq. No.
Contig ID
                   4327_3.R1040
5'-most EST
                   g4305006
                   BLASTX
Method
                   g2688824
NCBI GI
BLAST score
                   283
E value
                   3.0e-31
Match length
                   121
```

NCBI Description

% identity

on (U93273) putative auxin-repressed protein [Prunus

armeniaca]

5019

Seq. No.



```
4327 4.R1040
Contig ID
5'-most EST
                  rca700996376.hl
Seq. No.
                  5020 -
Contig ID
                  4327 5.R1040
5'-most EST
                  q4295801
Method
                  BLASTX
                  q2688824
NCBI GI
BLAST score
                  367
                  6.0e-35
E value
Match length
                  92
                  78
% identity
                  (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                  armeniaca]
Seq. No.
                  5021
                  4327 6.R1040
Contig ID
5'-most EST
                  pxt700943152.h1
Method
                  BLASTX
                  q2688824
NCBI GI
BLAST score
                  226
E value
                  2.0e-18
Match length
                  61
% identity
                  79
NCBI Description
                  (U93273) putative auxin-repressed protein [Prunus
                  armeniaca]
Seq. No.
                  5022
Contig ID
                  4330 1.R1040
5'-most EST
                  LIB3049-005-Q1-E1-B8
Method
                  BLASTX
                  q2791806
NCBI GI
BLAST score
                  523
                  5.0e-53
E value
                  165
Match length
% identity
                  58
                  (AF041433) bet3 [Mus musculus]
NCBI Description
Seq. No.
                  5023
Contig ID
                  4331 1.R1040
5'-most EST
                  LIB3138-020-Q1-N1-A9
Method
                  BLASTX
NCBI GI
                  g731672
BLAST score
                  318
                  2.0e-29
E value
Match length
                  124
% identity
                  48
                  HYPOTHETICAL 56.5 KD PROTEIN IN DYS1-ERG7 INTERGENIC REGION
NCBI Description
                  >gi_626628_pir__S46697 hypothetical protein YHR070w - yeast
                   (Saccharomyces cerevisiae) >gi_487945 (U00061) Yhr070wp
                   [Saccharomyces cerevisiae]
                  5024
Seq. No.
```

Contig ID 4335_1.R1040 5'-most EST ncj700977406.h1

Method BLASTN NCBI GI g2330565



```
BLAST score
                   5.0e-11
E value
                  141
Match length
                   86
% identity
NCBI Description Arabidopsis thaliana rpoMt gene
                  5025
Seq. No.
                   4335 2.R1040
Contig ID
5'-most EST
                  zhf700953673.h1
                   5026
Seq. No.
                   4340 1.R1040
Contig ID
                   leu701150583.h1
5'-most EST
                  BLASTX
Method
                   g3559805
NCBI GI
BLAST score
                   363
E value
                   2.0e-34
                   79
Match length
                   80
% identity
NCBI Description (AJ006787) putative phytochelatin synthetase [Arabidopsis
                   thaliana]
Seq. No.
                   5027
                   4340 2.R1040
Contig ID
                   zzp700834154.h1
5'-most EST
                   5028
Seq. No.
Contig ID
                   4346 1.R1040
                   hrw7\overline{0}1059215.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1657855
BLAST score
                   502
E value
                   2.0e-50
                   201
Match length
% identity
                   (U73216) cold acclimation protein WCOR413 [Triticum
NCBI Description
                   aestivum]
                   5029
Seq. No.
Contig ID
                   4346 2.R1040
5'-most EST
                   q5607172
Seq. No.
                   5030
                   4346 3.R1040
Contig ID
                   LIB3170-086-Q1-K1-C7
5'-most EST
Method
                   BLASTX
                   g1657855
NCBI GI
BLAST score
                   221
                   1.0e-17
E value
Match length
                   54
                   70
% identity
NCBI Description
                   (U73216) cold acclimation protein WCOR413 [Triticum
                   aestivum]
```

Seq. No. Contig ID 5031

4346 4.R1040

5'-most EST

LIB3093-048-Q1-K1-F5



```
Method
                  BLASTX
NCBI GI
                  g1657855
BLAST score
                  129
                  2.0e-20
E value
                  80
Match length
                   66
% identity
                  (U73216) cold acclimation protein WCOR413 [Triticum
NCBI Description
                  aestivum]
                  5032
Seq. No.
                  4347 2.R1040
Contig ID
                  LIB3028-017-Q1-B1-G7
5'-most EST
                  5033
Seq. No.
                  4348 1.R1040
Contig ID
                  LIB3049-046-Q1-E1-A3
5'-most EST
Method
                  BLASTN
                  g3821292
NCBI GI
BLAST score
                  65
                  7.0e-28
E value
                  196
Match length
                  84
% identity
NCBI Description Pisum sativum cDNA fragment generated by DDRT-Alu
                   5034
Seq. No.
                   4351 1.R1040
Contig ID
5'-most EST
                  LIB3040-059-Q1-E1-E6
Method
                  BLASTN
NCBI GI
                   g2565416
BLAST score
                   147
                   2.0e-76
E value
                   511
Match length
% identity
                   83
                  Onobrychis viciifolia oxygen-evolving enhancer protein 3
NCBI Description
                   precursor, mRNA, complete cds
Seq. No.
                   5035
Contig ID
                   4351 2.R1040
                   asn701133531.h2
5'-most EST
Method
                   BLASTN
NCBI GI
                   g2565416
BLAST score
                   42
                   2.0e-14
E value
Match length
                   126
% identity
                   83
                  Onobrychis viciifolia oxygen-evolving enhancer protein 3
NCBI Description
                   precursor, mRNA, complete cds
                   5036
Seq. No.
                   4353 1.R1040
Contig ID
                   jC-gmle01810044d07d1
5'-most EST
Method
                   BLASTX
```

Method BLASTX
NCBI GI g2262165
BLAST score 436
E value 6.0e-43
Match length 98
% identity 87



```
(AC002329) predicted protein of unknown function
NCBI Description
                   [Arabidopsis thaliana]
                  5037
Seq. No.
Contig ID
                  4353 2.R1040
5'-most EST
                  LIB3051-003-Q1-E1-D9
Method
                  BLASTX
NCBI GI
                  q2262165
BLAST score
                  528
                  2.0e-55
E value
Match length
                  174
                  70
% identity
                   (AC002329) predicted protein of unknown function
NCBI Description
                   [Arabidopsis thaliana]
                   5038
Seq. No.
Contig ID
                   4354 1.R1040
5'-most EST
                   jex700908542.h1
Method
                  BLASTX
                  q4099092
NCBI GI
                   475
BLAST score
E value
                  8.0e-48
                  115
Match length
                   77
% identity
                  (U83179) unknown [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5039
Contig ID
                   4355 1.R1040
5'-most EST
                  uC-gmrominsoy202c05b1
Method
                   BLASTX
NCBI GI
                   g2924792
BLAST score
                   417
E value
                   1.0e-40
Match length
                   112
% identity
                   71
                   (AC002334) similar to synaptobrevin [Arabidopsis thaliana]
NCBI Description
                   5040
Seq. No.
                   4355 2.R1040
Contig ID
                   uC-gmrominsoy135e12b1
5'-most EST
Method
                   BLASTX
                            ist.
NCBI GI
                   q2914706
BLAST score
                   863
                   8.0e-93
E value
                   195
Match length
                   81
% identity
                   (AC003974) putative homeobox protein [Arabidopsis thaliana]
NCBI Description
                   5041
Seq. No.
Contig ID
                   4355 3.R1040
5'-most EST
                   pxt700944245.hl
Method
                   BLASTX
```

Method BLASTX
NCBI GI g2914706
BLAST score 336
E value 2.0e-31

Match length 78 % identity 74

```
(AC003974) putative homeobox protein [Arabidopsis thaliana]
 NCBI Description
                    5042
 Seq. No.
 Contig ID
                    4355 5.R1040
 5'-most EST
                    jC-gmro02910027h02d1
 Method
                    BLASTX
                    q1174498
 NCBI GI
 BLAST score
                    303
                    1.0e-27
 E value
 Match length
                    91
 % identity
                    SYNAPTOBREVIN-RELATED PROTEIN >gi_600710 (M90418) formerly
 NCBI Description
                    called HAT24; synaptobrevin-related protein [Arabidopsis
                    thaliana]
 Seq. No.
                    5043
 Contig ID
                    4355 6.R1040
                    uC-gmrominsoy170g03b1
5'-most EST
 Method
                    BLASTX
 NCBI GI
                    q2914706
 BLAST score
                    430
 E value
                    2.0e-42
 Match length
                    98
                    79
 % identity
                    (AC003974) putative homeobox protein [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
                    5044
 Contig ID
                    4357 1.R1040
 5'-most EST
                    rca701001244.h1
                    5045
 Seq. No.
 Contig ID
                    4358 1.R1040
                    k117\overline{0}1205513.h1
  5'-most EST
```

Method BLASTN
NCBI GI g2677829
BLAST score 198
E value 1.0e-107
Match length 498
% identity 85

NCBI Description Prunus armeniaca ribosomal protein L12 mRNA, complete cds

Seq. No. 5046

Contig ID 4358 2.R1040

5'-most EST LIB3093-028-Q1-K1-E5

Method BLASTN
NCBI GI g2677829
BLAST score 181
E value 5.0e-97
Match length 505
% identity 84

NCBI Description Prunus armeniaca ribosomal protein L12 mRNA, complete cds

Seq. No. 5047

Contig ID 4358_3.R1040

5'-most EST LIB3073-004-Q1-K1-B11

Method BLASTX NCBI GI g2677830



```
BLAST score
                  181
E value
                  3.0e-13
Match length
                  37
% identity
                  89
                  (U93168) ribosomal protein L12 [Prunus armeniaca]
NCBI Description
                  5048
Seq. No.
                  4361 1.R1040
Contig ID
5'-most EST
                  txt700735883.h1
Method
                  BLASTN
                  g1617035
NCBI GI
BLAST score
                  250
                  1.0e-138
E value
Match length
                  385
% identity
                  92
NCBI Description V.unguiculata mRNA for Ted2 protein
                  5049
Seq. No.
                   4363 1.R1040
Contig ID
5'-most EST
                  rca700997990.hl
                  5050
Seq. No.
                   4368 1.R1040
Contig ID
                   sat701007339.h2
5'-most EST
                  BLASTX
Method
NCBI GI
                   g1724114
BLAST score
                   320
E value
                   6.0e-29
Match length
                  285
% identity
                   41
                  (U80041) Af10-protein [Avena fatua]
NCBI Description
                   5051
Seq. No.
                   4368 2.R1040
Contig ID
5'-most EST
                   jC-gmf102220073a05d1
                   5052
Seq. No.
Contig ID
                   4369 1.R1040
5'-most EST
                   dpv701097973.h1
Method
                   BLASTX
NCBI GI
                   g1407705
BLAST score
                   533
E value
                   1.0e-54
Match length
                   150
% identity
                   67
                  (U60202) lipoxygenase [Solanum tuberosum]
NCBI Description
Seq. No.
                   5053
Contig ID
                   4370 1.R1040
5'-most EST
                   uC-gmrominsoy029f01b1
Method
                   BLASTX
NCBI GI
                   q1899050
BLAST score
                   464
```

E value 2.0e-46

Match length 141 % identity

NCBI Description (U62773) ripening-associated protein [Lycopersicon

BLAST score

Match length

E value

464

162

4.0e-46



esculentum]

Seq. No. 5054 Contig ID 4370 2.R1040 5'-most EST LIB3056-012-Q1-N1-B7 Method BLASTX NCBI GI q1899050 972 BLAST score E value 1.0e-105 Match length 256 % identity 72 (U62773) ripening-associated protein [Lycopersicon NCBI Description esculentum] Seq. No. 5055 4371 1.R1040 Contig ID 5'-most EST LIB3049-004-Q1-E1-D10 5056 Seq. No. Contig ID 4374 1.R1040 5'-most EST uC-gmropic028a08b1 5057 Seq. No. Contig ID 4375 1.R1040 5'-most EST jC-gmf102220137c08a1 Method BLASTX NCBI GI g99809 1209 BLAST score E value 1.0e-133 Match length 326 % identity 71 NCBI Description gene Bp10 protein - rape >gi 17795 emb CAA47177 (X66608) Bplo [Brassica napus] 5058 Seq. No. Contig ID 4376 1.R1040 5'-most EST LIB3028-017-Q1-B1-D1 Method BLASTX NCBI GI g2642157 BLAST score 223 E value 3.0e-18 Match length 68 57 % identity NCBI Description (AC003000) ankyrin-like protein [Arabidopsis thaliana] 5059 Seq. No. 4377 1.R1040 Contig ID 5'-most EST kmv700742027.h1 5060 Seq. No. Contig ID 4379 1.R1040 5'-most EST uC-gmrominsoy208f01b1 Method BLASTX NCBI GI g4455250



```
% identity
                  (AL035523) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                  4379 2.R1040
5'-most EST
                  LIB3051-026-Q1-K1-F4
                  BLASTX
Method
                  q4455250
NCBI GI
BLAST score
                  394
E value
                  3.0e-38
Match length
                  121
                  50
% identity
                  (AL035523) putative protein [Arabidopsis thaliana]
NCBI Description
                  5062
Seq. No.
Contig ID
                  4385 1.R1040
                  ssr700559413.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q433608
BLAST score
                  268
                  1.0e-148
E value
Match length
                  851
                  83
% identity
NCBI Description
                  R.communis mRNA for enolase
                   5063
Seq. No.
Contig ID
                   4385 2.R1040
5'-most EST
                  LIB3028-009-Q1-B1-B2
Method
                  BLASTN
NCBI GI
                   g1041244
BLAST score
                   41
                   1.0e-13
E value
Match length
                  89
                  87
% identity
                  A.glutinosa mRNA for enolase
NCBI Description
                   5064
Seq. No.
                   4385 3.R1040
Contig ID
5'-most EST
                   zsg701118769.hl
                  BLASTX
Method
NCBI GI
                   g2673918
                   197
BLAST score
                   6.0e-15
E value
Match length
                   150
% identity
                   37
NCBI Description
                  (AC002561) unknwon protein [Arabidopsis thaliana]
                   5065
Seq. No.
                   4389 1.R1040
Contig ID
                   asn701143195.hl
Method
                   BLASTX
```

5'-most EST

NCBI GI q4522008 463 BLAST score 2.0e-46 E value Match length 149 % identity 59

NCBI Description (AC007069) hypothetical protein [Arabidopsis thaliana]

E value

Match length



```
5066
Seq. No.
                  4392 1.R1040
Contig ID
5'-most EST
                  LIB3170-019-Q1-K1-G2
Method
                  BLASTX
                  g3046693
NCBI GI
BLAST score
                  722
                  2.0e-76
E value
                  137
Match length
% identity
                  92
                  (AL022140) receptor like protein (fragment) [Arabidopsis
NCBI Description
                  thaliana]
                   5067
Seq. No.
Contig ID
                   4392 2.R1040
                  jC-gmle01810006h04a1
5'-most EST
Method
                  BLASTX
NCBI GI
                   q3046693
BLAST score
                   220
                   7.0e-18
E value
Match length
                   50
% identity
                   80
                   (AL022140) receptor like protein (fragment) [Arabidopsis
NCBI Description
                   thaliana]
                   5068
Seq. No.
Contig ID
                   4392 3.R1040
5'-most EST
                   LIB3049-019-Q1-E1-A11
Method
                   BLASTX
                   q3046693
NCBI GI
                   150
BLAST score
E value
                   8.0e-10
Match length
                   32
                   84
% identity
                   (AL022140) receptor like protein (fragment) [Arabidopsis
NCBI Description
                   thaliana]
                   5069
Seq. No.
                   4396 1.R1040
Contig ID
                   LIB3138-033-Q1-N1-D5
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3360289
BLAST score
                   1017
                   1.0e-146
E value
                   373
Match length
                   67
% identity
                   (AF023164) leucine-rich repeat transmembrane protein kinase
NCBI Description
                   1 [Zea mays]
                   5070
Seq. No.
Contig ID
                   4396 2.R1040
                   fC-qmle700685276b1
5'-most EST
                   BLASTX
Method
                   q3360291
NCBI GI
                   313
BLAST score
```

1.0e-28 81



```
% identity
                  (AF023165) leucine-rich repeat transmembrane protein kinase
NCBI Description
                  2 [Zea mays]
                  5071
Seq. No.
Contig ID
                  4398 1.R1040
                  dpv701099682.hl
5'-most EST
Method
                  BLASTX
                  a544075
NCBI GI
                  563
BLAST score
                  6.0e-58
E value
Match length
                  135
                  28
% identity
                  COATOMER BETA' SUBUNIT (BETA'-COAT PROTEIN) (BETA'-COP)
NCBI Description
                  (P102) >gi 486768 pir S35312 coatomer complex chain beta'
                  - bovine >gi_312732_emb_CAA51285_ (X72756) beta prime cop
                  [Bos taurus]
                  5072
Seq. No.
                  4398 2.R1040
Contig ID
                  uC-qmropic074d04b1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q544075
                  540
BLAST score
                  3.0e-55
E value
                  129
Match length
% identity
                  COATOMER BETA' SUBUNIT (BETA'-COAT PROTEIN) (BETA'-COP)
NCBI Description
                  (P102) >gi 486768 pir _S35312 coatomer complex chain beta'
                  - bovine >gi_312732 emb_CAA51285_ (X72756) beta prime cop
                   [Bos taurus]
                  5073
Seq. No.
                  4398 3.R1040
Contig ID
                  jex700905015.h1
5'-most EST
Method
                  BLASTX
                  g3023522
NCBI GI
                  2067
BLAST score
                  0.0e + 00
E value
                   660
Match length
% identity
                   53
                  COATOMER BETA' SUBUNIT (BETA'-COAT PROTEIN) (BETA'-COP)
NCBI Description
                   (P102) >qi 2454309 (AF002705) beta prime COP [Rattus
                  norvegicus]
Seq. No.
                   5074
                   4398 4.R1040
Contig ID
```

fua701037018.h1 5'-most EST

BLASTX Method q544075 NCBI GI BLAST score 503 6.0e-51 E value 130 Match length % identity

COATOMER BETA' SUBUNIT (BETA'-COAT PROTEIN) (BETA'-COP) NCBI Description (P102) >gi_486768_pir__S35312 coatomer complex chain beta' - bovine >gi_312732_emb_CAA51285_ (X72756) beta prime cop



```
[Bos taurus]
                   5075
Seq. No.
                   4398 5.R1040
Contig ID
5'-most EST
                   jC-gmst02400059b04a1
Method
                  BLASTN
                  q3046855
NCBI GI
                   36
BLAST score
                   1.0e-10
E value
                   96
Match length
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:
NCBI Description
                  MSL1, complete sequence [Arabidopsis thaliana]
                   5076
Seq. No.
Contig ID
                   4400 1.R1040
                   LIB3028-017-Q1-B1-C12
5'-most EST
Method
                   BLASTX
                   q3885329
NCBI GI
                   458
BLAST score
                   5.0e-53
E value
Match length
                   131
% identity
                   (AC005623) alien-like protein [Arabidopsis thaliana]
NCBI Description
                   5077
Seq. No.
Contig ID
                   4408 1.R1040
                   jC-qmst02400060g09a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q417256
                   159
BLAST score
E value
                   6.0e-10
                   164
Match length
% identity
                   32
```

LIPASE PRECURSOR (TRIACYLGLYCEROL LIPASE) NCBI Description

5078 Seq. No. 4408 3.R1040 Contig ID 5'-most EST ncj700987853.hl

5079 Seq. No. 4410 1.R1040 Contig ID taw700657037.h1 5'-most EST

BLASTX Method g140739 NCBI GI BLAST score 283 6.0e-25 E value 169 Match length 36 % identity

HYPOTHETICAL 31.0 KD PROTEIN IN RNPB-SOHA INTERGENIC REGION NCBI Description

(ORF 2) >gi_78519_pir__JQ0613 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) homolog - Escherichia coli >gi_216632_dbj_BAA14238_ (D90212) ORF2 [Escherichia coli]
>gi_606065 (U18997) ORF_f299 [Escherichia coli] >gi_1789513

(AE000394) putative dehydrogenase [Escherichia coli]

Seq. No. 5080



```
Contig ID
                  4411 1.R1040
5'-most EST
                  LIB3028-017-Q1-B1-A2
Seq. No.
                  5081
Contig ID
                  4413 1.R1040
                  asn701132849.h1
5'-most EST
Method
                  BLASTX
                  q4567255
NCBI GI
BLAST score
                  512
                  5.0e-56
E value
Match length
                  183
% identity
                  66
                  (AC007070) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  5082
Contig ID
                  4413 2.R1040
5'-most EST
                  LIB3106-115-Q1-K1-H2
Method
                  BLASTN
                  g1177368
NCBI GI
BLAST score
                  86
                  2.0e-40
E value
Match length
                  158
% identity
                  89
                  Z.mays mRNA for ribosomal protein L39
NCBI Description
Seq. No.
                  5083
                  4413_3.R1040
Contig ID
5'-most EST
                  jC-gmst02400053c10d1
Method
                  BLASTN
NCBI GI
                  q1177368
BLAST score
                  86
E value
                  1.0e-40
Match length
                  158
% identity
                  89
NCBI Description Z.mays mRNA for ribosomal protein L39
                  5084
Seq. No.
Contig ID
                  4413 5.R1040
5'-most EST
                  g4292679
Method
                  BLASTN
NCBI GI
                  g1177368
BLAST score
                  67
E value
                  3.0e-29
Match length
                  159
% identity
                  86
NCBI Description Z.mays mRNA for ribosomal protein L39
Seq. No.
                  5085
                  4415 1.R1040
Contig ID
5'-most EST
                  LIB3109-048-Q1-K1-C6
Method
                  BLASTX
NCBI GI
                  g1170606
BLAST score
                  302
E value
                  7.0e-34
Match length
                  120
% identity
                  59
```

NCBI Description ADENYLATE KINASE, CHLOROPLAST (ATP-AMP TRANSPHOSPHORYLASE)



5086

42

>gi_629863_pir__S45634 adenylate kinase (EC 2.7.4.3),
chloroplast - maize >gi_3114421_pdb_1ZAK_A Chain A,
Adenylate Kinase From Maize In Complex With The Inhibitor
P1,P5-Bis(Adenosine-5'-)pentaphosphate (Ap5a)
>gi_3114422_pdb_1ZAK_B Chain B, Adenylate Kinase From Maize
In Complex With The Inhibitor
P1,P5-Bis(Adenosine-5'-)pentaphosphate (Ap5a)

Contig ID 4416 1.R1040
5'-most EST kmv700738152.h1
Method BLASTX
NCBI GI g2505879
BLAST score 373
E value 2.0e-35
Match length 251

NCBI Description (Y12776) hypothetical protein [Arabidopsis thaliana]

Seq. No. 5087

Seq. No.

% identity

Contig ID 4416_2.R1040 5'-most EST leu701147855.h1

Seq. No. 5088

Contig ID 4416_3.R1040 5'-most EST uxk700669004.h1

Seq. No. 5089

Contig ID 4418 1.R1040

5'-most EST LIB3051-114-Q1-K1-A1

Method BLASTX
NCBI GI g136140
BLAST score 374
E value 2.0e-35
Match length 195
% identity 38

NCBI Description PUTATIVE AC9 TRANSPOSASE >gi_72973_pir__TQZMCA probable

transposase - maize transposon Ac9

Seq. No. 5090

Contig ID 4419 1.R1040

5'-most EST LIB3\overline{109}-022-Q1-K1-G4

Method BLASTX
NCBI GI g2760325
BLAST score 671
E value 1.0e-134
Match length 360
% identity 66

NCBI Description (AC002130) F1N21.10 [Arabidopsis thaliana]

Seq. No. 5091

Contig ID 4419_2.R1040

5'-most EST LIB3139-074-P1-N1-B5

Method BLASTX NCBI GI g2760325 BLAST score 249 E value 2.0e-21



```
Match length
                  67
% identity
                  72
                  (ACO02130) F1N21.10 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  5092
Contig ID
                  4419 3.R1040
5'-most EST
                  uC-gmrominsoy316c12b1
Method
                  BLASTX
NCBI GI
                  q2760325
BLAST score
                  366
E value
                  8.0e-51
Match length
                  137
% identity
                  77
NCBI Description
                  (AC002130) F1N21.10 [Arabidopsis thaliana]
Seq. No.
                  5093
Contig ID
                  4420 1.R1040
5'-most EST
                  LIB3107-002-Q1-K1-E9
Method
                  BLASTX
NCBI GI
                  q3218467
BLAST score
                  154
E value
                  7.0e-10
Match length
                  149
                  31
% identity
NCBI Description (AJ006529) putative phosphatase [Gallus gallus]
Seq. No.
                  5094
Contig ID
                  4422 1.R1040
5'-most EST
                  zhf700957434.hl
Method
                  BLASTN
NCBI GI
                  q166379
BLAST score
                  565
E value
                  0.0e+00
Match length
                  1241
% identity
                  83
                  Alfalfa glucose-regulated endoplasmic reticular protein
NCBI Description
                  mRNA, complete cds
                  5095
Seq. No.
                  4422 2.R1040
Contig ID
5'-most EST
                  kmv700739677.h1
Method
                  BLASTN
                  g166379
NCBI GI
BLAST score
                  66
E value
                  1.0e-28
                  180
Match length
% identity
                  84
                  Alfalfa glucose-regulated endoplasmic reticular protein
NCBI Description
                  mRNA, complete cds
                  5096
```

Seq. No.

Contig ID 4422 3.R1040

5'-most EST LIB3051-010-Q1-E1-E7

Method BLASTN NCBI GI g166379 BLAST score 83 8.0e-39 E value



```
Match length
                  164
                  92
% identity
                  Alfalfa glucose-regulated endoplasmic reticular protein
NCBI Description
                  mRNA, complete cds
                  5097
Seq. No.
                  4422 4.R1040
Contig ID
                  jC-gmle01810070g05a1
5'-most EST
                  BLASTN
Method
                  g166379
NCBI GI
                  43
BLAST score
                  9.0e-15
E value
                  164
Match length
                  87
% identity
                  Alfalfa glucose-regulated endoplasmic reticular protein
NCBI Description
                  mRNA, complete cds
                   5098
Seq. No.
                   4427 1.R1040
Contig ID
                   jC-gmst02400058c09d1
5'-most EST
                   BLASTX
Method
                   g4033735
NCBI GI
BLAST score
                   209
                   6.0e-29
E value
                   86
Match length
% identity
                  (AF054284) spliceosomal protein SAP 155 [Homo sapiens]
NCBI Description
Seq. No.
                   5099
                   4433 1.R1040
Contig ID
                   LIB3139-071-P1-N1-C7
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2129473
                   282
BLAST score
                   1.0e-24
E value
                   145
Match length
                   43
% identity
                   arabinogalactan-like protein - loblolly pine >gi_607774
NCBI Description
                   (U09556) arabinogalactan-like protein [Pinus taeda]
Seq. No.
                   5100
                   4433 2.R1040
Contig ID
5'-most EST
                   jC-gmst02400002c01d1
                   5101
Seq. No.
                   4436 1.R1040
Contig ID
                   LIB3028-016-Q1-B1-G9
5'-most EST
Method
                   BLASTX
                   g3415113
NCBI GI
                   300
BLAST score
                   1.0e-34
E value
Match length
                   133
```

Seq. No. 5102

% identity

NCBI Description

Contig ID 4438_1.R1040

39

(AF081201) villin 1 [Arabidopsis thaliana]



```
sat701003234.hl
5'-most EST
Method
                  BLASTX
                  g401621
NCBI GI
BLAST score
                  245
                  1.0e-20
E value
Match length
                  137
% identity
                  42
                  HYPOTHETICAL 20.4 KD PROTEIN IN TNAB-BGLB INTERGENIC REGION
NCBI Description
                  >gi 290561 (L10328) o188 [Escherichia coli] >gi_1790149
                   (AE000448) orf, hypothetical protein [Escherichia coli]
                  5103
Seq. No.
                  4440 1.R1040
Contig ID
                  LIB3039-021-Q1-E1-E4
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3650368
BLAST score
                  687
                  3.0e-72
E value
                  190
Match length
                   69
% identity
                  (AJ011398) profucosidase [Pisum sativum]
NCBI Description
Seq. No.
                   5104
                   4442 1.R1040
Contig ID
                   zzp700834319.h1
5'-most EST
                   5105
Seq. No.
                   4447 1.R1040
Contig ID
                   uC-gmropic009a06b1
5'-most EST
Method
                   BLASTX
                   q3522950
NCBI GI
BLAST score
                   354
E value
                   2.0e-33
                   115
Match length
% identity
                   (AC004411) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5106
                   4448 1.R1040
Contig ID
                   leu701157807.hl
5'-most EST
Method
                   BLASTX
                   q3402282
NCBI GI
BLAST score
                   442
                   1.0e-43
E value
                   155
Match length
                   55
% identity
                  (AJ000997) proline-rich protein [Solanum tuberosum]
NCBI Description
                   5107
Seq. No.
                   4450 1.R1040
Contig ID
```

1

g4302030 5'-most EST

Seq. No.

5108

4450 2.R1040 Contig ID

LIB3107-055-Q1-K1-C11 5'-most EST

Method BLASTN NCBI GI g170067

BLAST score

37



```
BLAST score
                   540
                   0.0e + 00
E value
                   559
Match length
                   99
% identity
                  Soybean (G.max) proline-rich cell wall protein (SbPRP3)
NCBI Description
                  gene, complete cds
                   5109
Seq. No.
                   4450 3.R1040
Contig ID
                   uC-gmrominsoy086e07b1
5'-most EST
Method
                   BLASTX
                   g2760330
NCBI GI
BLAST score
                   541
                   3.0e-55
E value
                   139
Match length
% identity
                   (AC002130) F1N21.15 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5110
                   4452 1.R1040
Contig ID
                   jC-gmle01810061c04a1
5'-most EST
Method
                   BLASTX
                   q3292814
NCBI GI
BLAST score
                   446
                   5.0e-44
E value
                   101
Match length
% identity
                   (AL031018) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4453 1.R1040
Contig ID
                   uC-gmrominsoy204h12b1
5'-most EST
Method
                   BLASTX
                   q4239845
NCBI GI
BLAST score
                   695
                   7.0e-73
E value
Match length
                   336
% identity
                   46
                   (AB015855) transcription factor TEIL [Nicotiana tabacum]
NCBI Description
                   5112
Seq. No.
                   4453 3.R1040
Contig ID
                   fC-gmro700869769f3
5'-most EST
                   BLASTX
Method
                   g4239845
NCBI GI
BLAST score
                   468
                   8.0e-47
E value
                   171
Match length
                   53
% identity
                   (AB015855) transcription factor TEIL [Nicotiana tabacum]
NCBI Description
Seq. No.
                   5113
                   4455 1.R1040
Contig ID
                   LIB3028-016-Q1-B1-E6
5'-most EST
                   BLASTN
Method
                   g4165340
NCBI GI
```



```
5.0e-11
E value
Match length
                  170
                  86
% identity
                  Arabidopsis thaliana chromosome I BAC F11M15 genomic
NCBI Description
                  sequence, complete sequence [Arabidopsis thaliana]
Seq. No.
                  5114
                  4456 1.R1040
Contig ID
                  LIB3039-040-Q1-E1-H1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g417148
BLAST score
                  743
                  9.0e-79
E value
                  226
Match length
                  63
% identity
                  PROBABLE GLUTATHIONE S-TRANSFERASE (HEAT SHOCK PROTEIN 26A)
NCBI Description
                   (G2-4) >gi 99912 pir A33654 heat shock protein 26A -
                  soybean >gi 169981 (M20363) Gmhsp26-A [Glycine max]
Seq. No.
                  4458 1.R1040
Contig ID
5'-most EST
                  kmv700743063.h1
Method
                  BLASTX
                  g3482967
NCBI GI
BLAST score
                  785
                  1.0e-83
E value
                  185
Match length
% identity
                   (AL031369) Protein phosphatase 2C-like protein [Arabidopsis
NCBI Description
                  thaliana] >gi_4559345_gb AAD23006.1_AC006585_1 (AC006585)
                  protein phosphatase 2C [Arabidopsis thaliana]
Seq. No.
                   5116
                   4459 1.R1040
Contig ID
                  LIB3106-087-Q1-K1-D7
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2498732
BLAST score
                   603
E value
                   2.0e-62
                   169
Match length
% identity
                   68
                   PROBABLE NADP-DEPENDENT OXIDOREDUCTASE P2
NCBI Description
                   >gi_1362014_pir__S57612 zeta-crystallin homolog -
                   Arabidopsis thaliana >gi 886430 emb CAA89262 (Z49268)
                   zeta-crystallin homologue [Arabidopsis thaliana]
                   5117
Seq. No.
Contig ID
                   4459 2.R1040
                   uC-gmropic010d10b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2498732
                   1064
```

BLAST score 1.0e-116 E value 276 Match length

% identity 72

PROBABLE NADP-DEPENDENT OXIDOREDUCTASE P2 NCBI Description

>gi 1362014_pir S57612 zeta-crystallin homolog -



Arabidopsis thaliana >gi 886430 emb CAA89262 (Z49268) zeta-crystallin homologue [Arabidopsis thaliana]

Seq. No. 5118

4459 3.R1040 Contig ID

5'-most EST LIB3092-053-Q1-K1-B8

Method BLASTX NCBI GI g2498731 BLAST score 541 2.0e-55 E value 153 Match length 65 % identity

PROBABLE NADP-DEPENDENT OXIDOREDUCTASE P1 NCBI Description

> >gi 1362013 pir__S57611 zeta-crystallin homolog -Arabidopsis thaliana >gi 886428 emb CAA89838 (Z49768)

zeta-crystallin homologue [Arabidopsis thaliana]

5119 Seq. No.

4459 4.R1040 Contig ID

LIB3051-108-Q1-K1-H10 5'-most EST

Method BLASTX NCBI GI g2498731 BLAST score 692 6.0e-73 E value Match length 173 75 % identity

PROBABLE NADP-DEPENDENT OXIDOREDUCTASE P1 NCBI Description

>gi 1362013 pir S57611 zeta-crystallin homolog -Arabidopsis thaliana >gi 886428 emb CAA89838 (Z49768)

zeta-crystallin homologue [Arabidopsis thaliana]

5120 Seq. No.

4459 5.R1040 Contig ID

5'-most EST LIB3050-018-Q1-E1-A3

5121 Seq. No.

4459 6.R1040 Contig ID ssr700553955.h15'-most EST

Method BLASTX NCBI GI q2498731 BLAST score 535 1.0e-54 E value Match length 150 % identity 67

NCBI Description PROBABLE NADP-DEPENDENT OXIDOREDUCTASE P1

>gi_1362013_pir__S57611 zeta-crystallin homolog -Arabidopsis thaliana >gi_886428_emb_CAA89838_ (Z49768)

zeta-crystallin homologue [Arabidopsis thaliana]

5122 Seq. No.

Contig ID 4466 2.R1040 5'-most EST ssr700554843.h1

Method BLASTX g2905893 NCBI GI BLAST score 577 1.0e-59 E value Match length 173



```
% identity
NCBI Description (U66424) fimbrin-like protein [Arabidopsis thaliana]
                  5123
Seq. No.
                  4467 1.R1040
Contig ID
5'-most EST
                  uaw700661776.hl
                  5124
Seq. No.
                   4467 2.R1040
Contig ID
                  uC-gmflminsoy019c05b1
5'-most EST
                   5125
Seq. No.
                   4467 3.R1040
Contig ID
                  LIB3107-030-Q1-K1-B2
5'-most EST
                   5126
Seq. No.
                   4467 4.R1040
Contig ID
                  hyd700727362.h1
5'-most EST
                   5127
Seq. No.
                   4469 1.R1040
Contig ID
                   sat701009455.h1
5'-most EST
                   5128
Seq. No.
                   4471 1.R1040
Contig ID
                   LIB3028-016-Q1-B1-C10
5'-most EST
                   BLASTX
Method
                   g2815897
NCBI GI
                   177
BLAST score
                   9.0e-13
E value
Match length
                   64
% identity
NCBI Description
                   (AF043279) protein kinase 1 [Toxoplasma gondii]
                   >gi 4325070 gb AAD17245 (AF118099) PITSLRE-like protein
                   kinase [Toxoplasma gondii]
                   5129
Seq. No.
Contig ID
                   4476 1.R1040
                   uC-gmrominsoy093d12b1
5'-most EST
                   BLASTX
Method
                   q967125
NCBI GI
                   2482
BLAST score
E value
                   0.0e + 00
Match length
                   487
% identity
                   (U08140) calcium dependent protein kinase [Vigna radiata]
NCBI Description
Seq. No.
                   5130
                   4476 2.R1040
Contig ID
                   LIB3093-018-Q1-K1-C1
5'-most EST
Method
                   BLASTN
                   g967124
NCBI GI
BLAST score
                   465
```

NCBI Description Vigna radiata Rwilcz calcium dependent protein kinase

0.0e + 00

819

90

E value

Match length

% identity



(CDPK) mRNA, complete cds

```
5131
Seq. No.
                  4476 3.R1040
Contig ID
                  LIB3109-010-Q1-K1-B6
5'-most EST
                  BLASTN
Method
NCBI GI
                  q967124
                  101
BLAST score
                  2.0e-49
E value
                  137
Match length
                   93
% identity
                  Vigna radiata Rwilcz calcium dependent protein kinase
NCBI Description
                   (CDPK) mRNA, complete cds
                   5132
Seq. No.
                   4476 4.R1040
Contig ID
                   ssr700555278.hl
5'-most EST
                   5133
Seq. No.
                   4478 1.R1040
Contig ID
                   LIB3029-010-Q1-B1-B10
5'-most EST
                   5134
Seq. No.
                   4479 1.R1040
Contig ID
                   gsv701049440.hl
5'-most EST
                   BLASTN
Method
                   g3204100
NCBI GI
                   103
BLAST score
                   3.0e-50
E value
                   407
Match length
                   85
% identity
                   Cicer arietinum mRNA for hypothetical protein, clone Can107
NCBI Description
Seq. No.
                   4480 1.R1040
Contig ID
                   leu7\overline{01151508.h1}
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2739279
                   302
BLAST score
                   2.0e-27
E value
                   110
Match length
% identity
                   57
                   (AJ223177) short chain alcohol dehydrogenase [Nicotiana
NCBI Description
                   tabacum] >gi_2791348_emb_CAA11154_ (AJ223178) short chain
                   alcohol dehydrogenase [Nicotiana tabacum]
                   5136
Seq. No.
Contig ID
                   4486 1.R1040
                   zzp700832950.hl
5'-most EST
                   BLASTN
Method
                   g1556445
NCBI GI
                   132
BLAST score
```

Hordeum vulgare alpha tubulin (tubA) mRNA, complete cds

4.0e-68

276

87

E value

Match length

NCBI Description

% identity



```
Seq. No.
                  5137
                  4487 1.R1040
Contig ID
                  asn701133308.hl
5'-most EST
                  BLASTX
Method
                  q3193316
NCBI GI
BLAST score
                  1488
E value
                  1.0e-166
                  419
Match length
                  70
% identity
                  (AF069299) contains similarity to nucleotide sugar
NCBI Description
                  epimerases [Arabidopsis thaliana]
                   5138
Seq. No.
                   4487 2.R1040
Contig ID
                  trc700568419.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2583123
BLAST score
                   415
E value
                   2.0e-40
                   116
Match length
% identity
                  (AC002387) putative nucleotide sugar epimerase [Arabidopsis
NCBI Description
                   thaliana]
                   5139
Seq. No.
                   4487 3.R1040
Contig ID
5'-most EST
                   LIB3049-041-Q1-E1-E5
Method
                   BLASTX
                   q3193316
NCBI GI
BLAST score
                   242
                   5.0e-23
E value
                   151
Match length
% identity
                   48
                   (AF069299) contains similarity to nucleotide sugar
NCBI Description
                   epimerases [Arabidopsis thaliana]
Seq. No.
                   5140
                   4487 4.R1040
Contig ID
                   LIB3028-015-Q1-B1-H7
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2583106
                   38
BLAST score
                   6.0e-12
E value
                   70
Match length
                   90
% identity
                   Arabidopsis thaliana chromosome II BAC F4L23 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
                   5141
Seq. No.
                   4489 1.R1040
Contig ID
```

LIB3170-030-Q1-K1-G5 5'-most EST

Seq. No.

4495 1.R1040 Contig ID

 $jC-g\overline{m}st02400020dg11d1$ 5'-most EST

5142

BLASTX Method NCBI GI g4539321

```
BLAST score
                  328
                  2.0e-30
E value
                  88
Match length
                  69
% identity
NCBI Description (AL035679) putative protein [Arabidopsis thaliana]
Seq. No.
                  5143
                  4498 1.R1040
Contig ID
                  eep700866223.h1
5'-most EST
                  BLASTX
Method
                  q4204277
NCBI GI
BLAST score
                  698
                  3.0e-73
E value
                  295
Match length
% identity
NCBI Description (AC004146) Hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   4498 2.R1040
Contig ID
                  LIB3139-020-P1-N1-G11
5'-most EST
                   5145
Seq. No.
                   4498 3.R1040
Contig ID
                  uC-gmrominsoy231a03b1
5'-most EST
Method
                  {\tt BLASTX}
                   g4204277
NCBI GI
BLAST score
                   310
E value
                   3.0e-28
                   133
Match length
                   53
% identity
NCBI Description (AC004146) Hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   5146
                   4498 4.R1040
Contig ID
5'-most EST
                   ssr700556564.hl
                   BLASTX
Method
NCBI GI
                   g4204277
BLAST score
                   623
E value
                   1.0e-64
Match length
                   259
                   57
% identity
                  (AC004146) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5147
                   4498 5.R1040
Contig ID
                   uC-gmrominsoy097b09b1
5'-most EST
                   BLASTX
Method
                   g4204277
NCBI GI
BLAST score
                   410
                   3.0e-40
E value
                   82
Match length
% identity
                   91
                   (AC004146) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5148
Seq. No.
```

988

4498 6.R1040

uC-gmrominsoy215d03b1

Contig ID

5'-most EST

E value

Match length

% identity

137

78



```
BLASTX
Method
                  q4204277
NCBI GI
                  274
BLAST score
                   3.0e-24
E value
                   55
Match length
% identity
                  (AC004146) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5149
Seq. No.
                   4500 1.R1040
Contig ID
                   jC-gmst02400011g06a1
5'-most EST
                   5150
Seq. No.
                   4500 2.R1040
Contig ID
                   LIB3107-010-Q1-K1-E8
5'-most EST
                   5151
Seq. No.
                   4500 3.R1040
Contig ID
                   vwf700673739.h1
5'-most EST
                   5152
Seq. No.
                   4504 1.R1040
Contig ID
                   uC-gmropic021c06b1
5'-most EST
                   BLASTN
Method
                   q1237085
NCBI GI
                   480
BLAST score
                   0.0e + 00
E value
                   1016
Match length
                   87
% identity
                   P.sativum mRNA for ADP-glucose pyrophosphorylase (agpl1)
NCBI Description
                   >gi_1599261_gb_I18906_I18906 Sequence 1 from patent US
                   5498831 >gi_3972107_gb_AR014653_AR014653 Sequence 1 from
                   patent US
                   5153
Seq. No.
                   4504 2.R1040
Contig ID
                   rlr700896089.h1
5'-most EST
                   BLASTN
Method
                   g1237085
NCBI GI
                   68
BLAST score
                   9.0e-30
E value
Match length
                   96
 % identity
                   93
                   P.sativum mRNA for ADP-glucose pyrophosphorylase (agpl1)
NCBI Description
                   >gi 1599261_gb_I18906_I18906 Sequence 1 from patent US
                    5498831 >gi_3972107_gb_AR014653_AR014653 Sequence 1 from
                   patent US
 Seq. No.
                    5154
                    4504 3.R1040
 Contig ID
                    uaw700663791.hl
 5'-most EST
                    BLASTX
 Method
 NCBI GI
                    q2625086
                    333
 BLAST score
                    1.0e-55
```

NCBI GI BLAST score

1242



```
(AF030383) ADP-glucose pyrophosphorylase large subunit
NCBI Description
                  [Cucumis melo var. markuwa Markino]
                  5155
Seq. No.
                  4506 1.R1040
Contig ID
                  LIB3028-015-Q1-B1-G12
5'-most EST
                  BLASTX
Method
                  q4406777
NCBI GI
                  198
BLAST score
                  2.0e-15
E value
Match length
                  53
                  79
% identity
                  (AC006532) putative zinc-finger protein [Arabidopsis
NCBI Description
                  thaliana]
                  5156
Seq. No.
                  4507 1.R1040
Contig ID
5'-most EST
                  LIB3028-015-Q1-B1-F1
                  BLASTX
Method
                  q3128186
NCBI GI
                  159
BLAST score
                  1.0e-10
E value
                  49
Match length
                   69
% identity
NCBI Description (AC004521) hypothetical protein [Arabidopsis thaliana]
                   5157
Seq. No.
                   4508 1.R1040
Contig ID
                   LIB3028-015-Q1-B1-G2
5'-most EST
Method
                   BLASTX
                   q3287688
NCBI GI
                   160
BLAST score
E value
                   7.0e-11
                   52
Match length
% identity
                   (AC003979) Contains similarity to ycf37 gene product
NCBI Description
                   gb_1001425 from Synechocystis sp. genome gb_D63999. ESTs
                   gb_T43026, gb_R64902, gb_Z18169 and gb_N37374 come from
                   this gene. [Arabidopsis thaliana]
                   5158
Seq. No.
                   4511 1.R1040
Contig ID
                   uC-gmrominsoy156g06b1
5'-most EST
                   BLASTX
Method
                   g2335108
NCBI GI
                   520
BLAST score
                   5.0e-53
E value
                   142
Match length
% identity
                   (AC002339) putative isulinase [Arabidopsis thaliana]
NCBI Description
                   5159
Seq. No.
                   4513 1.R1040
Contig ID
                   LIB3107-070-Q1-K1-E11
5'-most EST
                   BLASTX
Method
                   g4204265
```

```
1.0e-137
E value
Match length
                  403
% identity
NCBI Description
                  (AC005223) 45643 [Arabidopsis thaliana]
Seq. No.
                  4513 2.R1040
Contig ID
                  LIB3107-033-Q1-K1-C12
5'-most EST
                  BLASTX
Method
                  g4204265
NCBI GI
BLAST score
                  289
                  2.0e-25
E value
                  250
Match length
                  33
% identity
NCBI Description (AC005223) 45643 [Arabidopsis thaliana]
Seq. No.
Contig ID
                   4513 3.R1040
                  LIB3049-001-Q1-E1-E6
5'-most EST
                   5162
Seq. No.
Contig ID
                   4514 1.R1040
5'-most EST
                   zhf700958333.h1
                  BLASTX
Method
                   q2398679
NCBI GI
BLAST score
                   2001
E value
                   0.0e + 00
                   408
Match length
                   93
% identity
                  (Y14797) 3-deoxy-D-arabino-heptulosonate 7-phosphate
NCBI Description
                   synthase [Morinda citrifolia]
Seq. No.
                   5163
```

4514 2.R1040 Contig ID

LIB3139-055-P1-N1-D12 5'-most EST

BLASTX Method g114193 NCBI GI BLAST score 496 5.0e-50 E value 167 Match length % identity 60

PHOSPHO-2-DEHYDRO-3-DEOXYHEPTONATE ALDOLASE 1 PRECURSOR NCBI Description

(PHOSPHO-2-KETO-3-DEOXYHEPTONATE ALDOLASE 1) (DAHP

SYNTHETASE 1) (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE

SYNTHASE 1) >gi_170225 (M64261)

3-deoxy-D-arabino-heptulosonate 7-phosphate synthase

[Nicotiana tabacum] >gi_228697_prf__1808327A

deoxyheptulosonate phosphate synthase [Nicotiana tabacum]

5164 Seq. No.

4516 1.R1040 Contig ID $k117\overline{0}1211336.h1$ 5'-most EST

Method BLASTX q3790100 NCBI GI 2440 BLAST score E value 0.0e + 00559 Match length



```
% identity
                  (AF095520) pyrophosphate-dependent phosphofructokinase beta
NCBI Description
                  subunit [Citrus X paradisi]
                  5165
Seq. No.
Contig ID
                  4516 2.R1040
                  zsq701129503.h1
5'-most EST
                  BLASTX
Method
                  q3618343
NCBI GI
BLAST score
                  655
E value
                   4.0e-68
                  379
Match length
                  37
% identity
                   (AB009398) 26S proteasome subunit p40.5 [Homo sapiens]
NCBI Description
                  >gi 4506223 ref NP 002808.1 pPSMD13 proteasome (prosome,
                  macropain) 26S subunit, non-ATPase,
                   5166
Seq. No.
                   4516 3.R1040
Contig ID
                   fC-qmro700747787a2
5'-most EST
                  BLASTX
Method
NCBI GI
                   q3790100
BLAST score
                   205
                   5.0e-16
E value
                   65
Match length
% identity
                   (AF095520) pyrophosphate-dependent phosphofructokinase beta
NCBI Description
                   subunit [Citrus X paradisi]
Seq. No.
                   5167
                   4516 4.R1040
Contig ID
5'-most EST
                   epx701110260.h1
Method
                   BLASTX
                   g3790100
NCBI GI
BLAST score
                   710
                   4.0e-75
E value
                   176
Match length
                   79
% identity
                   (AF095520) pyrophosphate-dependent phosphofructokinase beta
NCBI Description
                   subunit [Citrus X paradisi]
                   5168
Seq. No.
                   4516 5.R1040
Contig ID
                   LIB3170-062-Q1-K1-B8
5'-most EST
                   BLASTX
Method
                   g3157931
NCBI GI
BLAST score
                   322
                   1.0e-29
E value
                   64
Match length
                   94
% identity
```

NCBI Description (AC002131) Similar to pyrophosphate-dependent

phosphofuctokinase beta subunit gb_Z32850 from Ricinus communis. ESTs gb_N65773, gb_N64925 and gb_F15232 come

from this gene. [Arabidopsis thaliana]

Seq. No. 5169

Contig ID 4516_6.R1040



```
fC-gmro700747787d3
5'-most EST
Seq. No.
                  5170
Contig ID
                  4516 7.R1040
5'-most EST
                  LIB3170-033-Q1-K1-D11
Method
                  BLASTN
NCBI GI
                  q2832611
BLAST score
                  59
E value
                  3.0e-24
                  205
Match length
% identity
                  87
                  Arabidopsis thaliana DNA chromosome 4, BAC clone F13C5
NCBI Description
                   (ESSAII project)
                  5171
Seq. No.
Contig ID
                  4516 10.R1040
5'-most EST
                  LIB3170-033-Q1-J1-D11
Seq. No.
                  5172
Contig ID
                  4516 11.R1040
5'-most EST
                  gsv701056017.h1
Method
                  BLASTN
                  g3790099
NCBI GI
BLAST score
                  79
                  1.0e-36
E value
Match length
                  107
% identity
                   93
                  Citrus X paradisi pyrophosphate-dependent
NCBI Description
                  phosphofructokinase beta subunit (PPi-PFKb) mRNA, complete
Seq. No.
                   5173
                   4516 12.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy271d04b1
                  BLASTX
Method
NCBI GI
                   g3790100
BLAST score
                   278
                   6.0e-25
E value
Match length
                   71
                   75
% identity
                   (AF095520) pyrophosphate-dependent phosphofructokinase beta
NCBI Description
                   subunit [Citrus X paradisi]
                   5174
Seq. No.
                   4518 1.R1040
Contig ID
5'-most EST
                   kwa701015659.h1
Method
                  BLASTX
                   g4204265
NCBI GI
BLAST score
                   459
                   2.0e-91
E value
Match length
                   294
                   59
% identity
                  (AC005223) 45643 [Arabidopsis thaliana]
NCBI Description
```

Seq. No. Contig ID

5175

Contig ID 4520_1.R1040 5'-most EST leu701147564.h1

% identity

NCBI Description

49



```
Method
                   BLASTX
NCBI GI
                   g2191177
BLAST score
                   194
                   7.0e-15
E value
Match length
                   67
% identity
                   61
                   (AF007270) belongs to the SPOU family of rRNA methylases.
NCBI Description
                   [Arabidopsis thaliana]
                   5176
Seq. No.
Contig ID
                   4521 1.R1040
5'-most EST
                   leu7\overline{01155507.h1}
Method
                   BLASTX
NCBI GI
                   q445612
BLAST score
                   600
E value
                   4.0e-62
Match length
                   122
% identity
                   93
NCBI Description ribosomal protein S19 [Solanum tuberosum]
                   5177
Seq. No.
Contig ID
                   4521 4.R1040
5'-most EST
                   jC-gmro02910067a06d1
                   BLASTX
Method
NCBI GI
                   q445612
BLAST score
                   252
E value
                   1.0e-21
Match length
                   53
% identity
                   92
NCBI Description ribosomal protein S19 [Solanum tuberosum]
Seq. No.
                   5178
                   4530 1.R1040
Contig ID
5'-most EST
                   LIB3170-039-Q1-K2-B7
                   5179
Seq. No.
                   4530 2.R1040
Contig ID
5'-most EST
                   LIB3107-018-Q1-K1-C3
Method
                   BLASTX
NCBI GI
                   q4467158
BLAST score
                   168
E value
                   2.0e-11
                   127
Match length
                   50
% identity
                  (AL035540) putative protein [Arabidopsis thaliana]
NCBI Description
                   5180
Seq. No.
                   4530_3.R1040
Contig ID
5'-most EST
                   LIB3170-045-Q1-J1-G10
Method
                   BLASTX
                   g4467158
NCBI GI
BLAST score
                   646
                   3.0e-67
E value
                   351
Match length
```

(AL035540) putative protein [Arabidopsis thaliana]



```
Seq. No.
                  5181
                  4530 5.R1040
Contig ID
5'-most EST
                  awf700839714.h1
Method
                  BLASTX
NCBI GI
                  g4467158
BLAST score
                  147
E value
                  1.0e-09
Match length
                  69
% identity
                  (AL035540) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                   4531 1.R1040
5'-most EST
                  seb700650590.hl
                  BLASTX
Method
NCBI GI
                  q1354849
BLAST score
                  501
E value
                   1.0e-50
Match length
                   154
% identity
NCBI Description (U57350) epoxide hydrolase [Nicotiana tabacum]
                   5183
Seq. No.
Contig ID
                   4531 3.R1040
5'-most EST
                  bth700849189.h1
Method
                   BLASTX
NCBI GI
                   g1354849
BLAST score
                   194
E value
                   5.0e-15
Match length
                   62
% identity
                  (U57350) epoxide hydrolase [Nicotiana tabacum]
NCBI Description
                   5184
Seq. No.
                   4533 1.R1040
Contig ID
5'-most EST
                   wvk700686367.h1
                   BLASTX
Method
                   g4455256
NCBI GI
                   875
BLAST score
                   4.0e-94
E value
                   228
Match length
% identity
                   71
                   (AL035523) protein-methionine-S-oxide reductase
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   5185
                   4533 2.R1040
Contig ID
                   LIB3170-073-Q1-K1-D2
5'-most EST
Seq. No.
                   5186
Contig ID
                   4533 3.R1040
                   jC-gmle01810027e12d1
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g1709690
BLAST score 318
E value 3.0e-29
Match length 75



% identity 80

NCBI Description PEPTIDE METHIONINE SULFOXIDE REDUCTASE (PEPTIDE MET(O)

REDUCTASE) >qi 1076454 pir S55365 probable peptide

methionine sulfoxide reductase - rape

>gi 853739 emb CAA88538 (Z48619) peptide methionine

sulfoxide reductase [Brassica napus]

>qi 1143406 emb CAA63919 (X94225) methionine sulfoxide

reductase [Brassica napus]

Seq. No. 5187

Contig ID 4533 4.R1040

5'-most EST jC-gmf102220070e08a1

Seq. No. 5188

Contig ID 4533_6.R1040 5'-most EST pxt700943002.h1

Seq. No. 5189

Contig ID 4537 1.R1040

5'-most EST LIB3053-007-Q1-N1-G10

Method BLASTX
NCBI GI g1652057
BLAST score 187
E value 1.0e-13
Match length 116
% identity 35

NCBI Description (D90902) hypothetical protein [Synechocystis sp.]

Seq. No. 5190

Contig ID 4542_1.R1040 5'-most EST epx701106433.h1

Method BLASTX
NCBI GI g1717975
BLAST score 353
E value 4.0e-33
Match length 123
% identity 54

NCBI Description 14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG) >gi_1177435_emb_CAA64670_ (X95384) 14.5

kDa translational inhibitor protein, p14.5 [Homo sapiens]

Seq. No. 5191

Contig ID 4552 1.R1040

5'-most EST LIB3\overline{1}39-116-P1-N1-C3

Method BLASTX
NCBI GI g4454472
BLAST score 227
E value 2.0e-18
Match length 118
% identity 43

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No. 5192

Contig ID 4552 2.R1040 5'-most EST zzp700831916.h1

Method BLASTX NCBI GI g4454472



```
421
BLAST score
                  8.0e-41
E value
                  228
Match length
                  46
% identity
                  (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                  5193
Seq. No.
                  4552 3.R1040
Contig ID
                  jC-gmro02800034c03a1
5'-most EST
                  5194
Seq. No.
                   4552 4.R1040
Contig ID
                  LIB3049-033-Q1-E1-E5
5'-most EST
                   5195
Seq. No.
                   4552 5.R1040
Contig ID
                  LIB3139-025-P1-N1-H6
5'-most EST
Seq. No.
                   5196
Contig ID
                   4554 1.R1040
                   hyd700728918.h1
5'-most EST
Method
                   BLASTN
                   q2502086
NCBI GI
                   53
BLAST score
                   1.0e-20
E value
                   118
Match length
                   93
% identity
                   Vigna radiata adenosine triphosphatase mRNA, complete cds
NCBI Description
                   5197
Seq. No.
                   4554 2.R1040
Contig ID
                   g5057934
5'-most EST
                   BLASTN
Method
                   g2502086
NCBI GI
                   370
BLAST score
                   0.0e+00
E value
                   552
Match length
                   93
% identity
NCBI Description Vigna radiata adenosine triphosphatase mRNA, complete cds
                   5198
Seq. No.
                   4554 3.R1040
Contig ID
5'-most EST
                   eep700868421.hl
Method
                   BLASTN
NCBI GI
                   g2502086
BLAST score
                   298
E value
                   1.0e-167
Match length
                   382
% identity
                   95
                   Vigna radiata adenosine triphosphatase mRNA, complete cds
NCBI Description
```

Seq. No. 5199

4554_4.R1040 Contig ID

LIB3170-058-Q1-J1-F2 5'-most EST

BLASTN Method g2502086 NCBI GI 180 BLAST score



E value 1.0e-96 Match length 392 87

NCBI Description Vigna radiata adenosine triphosphatase mRNA, complete cds

Seq. No. 5200

Contig ID 4554_5.R1040

5'-most EST LIB3087-010-Q1-K1-H3

Method BLASTN
NCBI GI g2502086
BLAST score 274
E value 1.0e-153
Match length 346
% identity 95

NCBI Description Vigna radiata adenosine triphosphatase mRNA, complete cds

Seq. No. 5201

Contig ID 4556_1.R1040

5'-most EST LIB3051-016-Q1-E1-A3

Method BLASTX
NCBI GI g4097579
BLAST score 947
E value 1.0e-102
Match length 199

matcn lengtn 195
% identity 90

NCBI Description (U64922) NTGP1 [Nicotiana tabacum]

Seq. No. 5202

Contig ID 4558_1.R1040 5'-most EST pmv700891157.h1

Seq. No. 5203

Contig ID 4559_1.R1040

5'-most EST LIB3051-091-Q1-K1-H10

Method BLASTX
NCBI GI g2811066
BLAST score 337
E value 5.0e-31
Match length 125
% identity 50

NCBI Description PUTATIVE PROTEIN-TYROSINE PHOSPHATASE PTEN (MUTATED IN

MULTIPLE ADVANCED CANCERS 1) >gi_1916330 (U92437) MMAC1

[Mus musculus]

Seq. No. 5204

Contig ID 4560_1.R1040

5'-most EST uC-gmronoir008e01b1

Method BLASTX
NCBI GI g3335378
BLAST score 251
E value 2.0e-21
Match length 136
% identity 40

NCBI Description (AC003028) Myb-related transcription activator [Arabidopsis

thaliana]

Seq. No. 5205

% identity

NCBI Description



```
4560 2.R1040
Contig ID
                  kl1701214579.hl
5'-most EST
                  BLASTX
Method
                  g4309734
NCBI GI
                  163
BLAST score
                  4.0e-11
E value
Match length
                  63
% identity
                  51
                   (AC006439) putative 26S proteosome regulatory subunit 8
NCBI Description
                   [Arabidopsis thaliana]
                  5206
Seq. No.
                   4565 1.R1040
Contig ID
                  LIB3170-039-Q1-K2-D5
5'-most EST
                   5207
Seq. No.
                   4565 2.R1040
Contig ID
                   uC-gmrominsoy164h05b1
5'-most EST
Method
                  BLASTX
                   g422029
NCBI GI
BLAST score
                   831
                   8.0e-90
E value
                   239
Match length
                   73
% identity
                   transcription factor OBF3.2, ocs element-binding - maize
NCBI Description
                   >qi 297018 emb CAA48904 (X69152) ocs-element binding
                   factor 3.2 [Zea mays]
                   5208
Seq. No.
                   4565 3.R1040
Contig ID
                   jC-gmro02910031a03a1
5'-most EST
                   BLASTX
Method
                   q1076782
NCBI GI
                   717
BLAST score
                   1.0e-75
E value
                   277
Match length
                   57
% identity
                   transcription factor HBP-1b(c1) - wheat (fragment)
NCBI Description
Seq. No.
                   4565 4.R1040
Contig ID
5'-most EST
                   wrg700787787.h2
Seq. No.
                   4565 6.R1040
Contig ID
5'-most EST
                   LIB3051-012-Q1-E1-E5
                   5211
Seq. No.
Contig ID
                   4567 1.R1040
                   LIB3106-099-Q1-K1-G10
5'-most EST
Method
                   BLASTX
                   g2384671
NCBI GI
BLAST score
                   845
                   2.0e-97
E value
                   244
Match length
                   73
```

(AF012657) putative potassium transporter AtKT2p



[Arabidopsis thaliana]

Seq. No. 5212

Contig ID 4569_1.R1040

5'-most EST uC-gmrominsoy099c05b1

Method BLASTX
NCBI GI g1076755
BLAST score 302
E value 2.0e-27
Match length 140
% identity 46

NCBI Description protein kinase - rice >gi_450300 (L27821) protein kinase

[Oryza sativa]

Seq. No. 5213

Contig ID 4569_2.R1040 5'-most EST uC-gmropic075a12b1

Seq. No. 5214

Contig ID 4570_1.R1040

5'-most EST LIB3028-014-Q1-B1-G6

Seq. No. 5215

Contig ID 4572 1.R1040 5'-most EST kl1701214909.h1

Method BLASTX
NCBI GI g2673904
BLAST score 283
E value 5.0e-25
Match length 85
% identity 75

NCBI Description (AC002561) hypothetical protein [Arabidopsis thaliana]

Seq. No. 5216

Contig ID 4575_1.R1040

5'-most EST uC-gmropic066e04b1

Method BLASTN
NCBI GI g3264766
BLAST score 90
E value 9.0e-43
Match length 206
% identity 86

NCBI Description Prunus armeniaca AP2 domain containing protein (AP2DCP)

mRNA, partial cds

Seq. No. 5217

Contig ID 4575_3.R1040 5'-most EST hrw701058671.h1

Method BLASTN
NCBI GI g2443456
BLAST score 53
E value 7.0e-21
Match length 69
% identity 94

NCBI Description Oryza sativa ethylene responsive element binding protein

(Os-EREBP1) mRNA, complete cds



Seq. No. 5218

4575 4.R1040 Contig ID 5'-most EST g5753147 Method BLASTX NCBI GI g3122602 BLAST score 443 E value 1.0e-43 Match length 168 % identity 17

NCBI Description PEROXISOMAL TARGETING SIGNAL 2 RECEPTOR (PTS2 RECEPTOR)

(PEROXIN-7) >gi_1890657 (U69171) peroxisomal PTS2 receptor

[Mus musculus]

Seq. No. 5219

Contig ID 4577 1.R1040

5'-most EST LIB3170-022-Q1-K1-G11

Method BLASTN
NCBI GI g3202041
BLAST score 219
E value 1.0e-119
Match length 759
% identity 82

NCBI Description Mesembryanthemum crystallinum 26S proteasome regulatory

subunit S5A mRNA, complete cds

Seq. No. 5220

Contig ID 4577_2.R1040 5'-most EST kl1701213417.h1

Method BLASTX
NCBI GI g3550985
BLAST score 295
E value 8.0e-27
Match length 61
% identity 92

NCBI Description (AB010740) OsS5a [Oryza sativa]

Seq. No. 5221

Contig ID 4580_1.R1040

5'-most EST LIB3106-056-Q1-K1-H8

Seq. No. 5222

Contig ID 4583 1.R1040

5'-most EST LIB3138-010-Q1-N1-F12

Method BLASTX
NCBI GI g4324495
BLAST score 1730
E value 0.0e+00
Match length 404
% identity 85

NCBI Description (AF105221) glutamyl-tRNA reductase precursor [Glycine max]

Seq. No. 5223

Contig ID 4586_1.R1040

5'-most EST jC-qmst02400023g07a1

Method BLASTX
NCBI GI g3142296
BLAST score 253



E value 9.0e-22 Match length 75 % identity 63

NCBI Description (AC002411) Contains similarity to hypothetical

mitochondrial import receptor subunit gb_Z98597 from S. pombe. ESTs gb_T45575 and gb_Z26435 and gb_AA394576 come

from this gene. [Arabidopsis thaliana]

Seq. No. 5224

Contig ID 4587_1.R1040 5'-most EST uxk700668616.h1

Method BLASTX
NCBI GI g2497281
BLAST score 1069
E value 1.0e-117
Match length 386
% identity 55

NCBI Description TRANSLATION INITIATION FACTOR IF-2

>gi 1651769 dbj BAA16696 (D90900) initiation factor IF-2

[Synechocystis sp.]

Seq. No. 5225

Contig ID 4589 1.R1040

5'-most EST LIB3028-014-Q1-B1-F1

Method BLASTX
NCBI GI g2982301
BLAST score 509
E value 1.0e-51
Match length 169
% identity 59

NCBI Description (AF051235) YGL010w-like protein [Picea mariana]

Seq. No. 5226

Contig ID 4592 1.R1040

5'-most EST LIB3\overline{107}-013-\overline{Q1}-\text{K1}-\overline{D9}

Method BLASTX
NCBI GI g2894603
BLAST score 182
E value 3.0e-13
Match length 66
% identity 36

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 5227

Contig ID 4594_1.R1040 5'-most EST kl1701205740.h1

Method BLASTX
NCBI GI g4539310
BLAST score 536
E value 4.0e-54
Match length 309
% identity 42

NCBI Description (AL035679) putative protein [Arabidopsis thaliana]

Seq. No.

5228

Contig ID 4594_2.R1040

5'-most EST LIB3107-058-Q1-K1-D9



```
5229
Seq. No.
                  4594 3.R1040
Contig ID
                  LIB3109-052-Q1-K1-F8
5'-most EST
                  5230
Seq. No.
Contig ID
                  4594 4.R1040
                  jex700910063.h1
5'-most EST
                  5231
Seq. No.
                  4594 5.R1040
Contig ID
5'-most EST
                  jC-gmf102220068g03d1
                  5232
Seq. No.
                   4595 1.R1040
Contig ID
                   jC-gmf102220088f03a1
5'-most EST
```

Method BLASTX g4406801 NCBI GI BLAST score 314 E value 6.0e-29 92 Match length 61

% identity

(AC006304) unknown protein [Arabidopsis thaliana] NCBI Description

5233 Seq. No. 4596 1.R1040 Contig ID sat701014394.h1 5'-most EST Method BLASTX g1658193 NCBI GI

BLAST score 905 E value 1.0e-97 Match length 238 % identity 72

(U74319) obtusifoliol 14-alpha demethylase CYP51 [Sorghum NCBI Description

bicolor]

5234 Seq. No.

Contig ID 4596 2.R1040 5'-most EST jex700907751.h1

Method BLASTX NCBI GI g2262164 BLAST score 183 E value 2.0e-13 Match length 79 % identity

(AC002329) putative obtusifoliol 14-alpha demethylase NCBI Description

[Arabidopsis thaliana]

5235 Seq. No.

Contig ID 4598 1.R1040 5'-most EST $k117\overline{0}1205289.h1$

Method BLASTX NCBI GI g3912953 BLAST score 224 5.0e-18 E value Match length 126 % identity 41

1003



NCBI Description PUTATIVE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE DEAMINASE (ACC DEAMINASE)

Seq. No. 5236

Contig ID 4601_1.R1040

5'-most EST LIB3028-014-Q1-B1-D2

Method BLASTX
NCBI GI g1350680
BLAST score 186
E value 8.0e-25
Match length 133
% identity 57

NCBI Description 60S RIBOSOMAL PROTEIN L1

5237

Seq. No.

Contig ID 4607 1.R1040

5'-most EST uC-gmropic103h01b1

Method BLASTX

NCBI GI g4039155

BLAST score 178

E value 2.0e-12

Match length 90

% identity 41

NCBI Description (AF104258) putative copper-inducible 35.6 kDa protein

[Festuca rubra]

Seq. No. 5238

Contig ID 4611_1.R1040

5'-most EST LIB3028-014-Q1-B1-A8

Seq. No. 5239

Contig ID 4612_1.R1040

5'-most EST LIB3170-026-Q1-K1-C11

Method BLASTX
NCBI GI g1707981
BLAST score 414
E value 2.0e-40
Match length 102
% identity 75

NCBI Description GLUTAREDOXIN >gi 1076561_pir__S54825 glutaredoxin - castor

bean

Seq. No. 5240

Contig ID 4615_1.R1040 5'-most EST leu701149172.h1

Method BLASTX
NCBI GI g4163997
BLAST score 1927
E value 0.0e+00
Match length 476
% identity 72

NCBI Description (AF087483) alpha-xylosidase precursor [Arabidopsis

thaliana]

Seq. No. 5241

Contig ID 4618_1.R1040

5'-most EST uC-gmrominsoy308a10b1



```
      Seq. No.
      5242

      Contig ID
      4619 1.R1040

      5'-most EST
      gsv701046306.h1

      Method
      BLASTX
```

NCBI GI g1498053
BLAST score 775
E value 2.0e-82
Match length 208
% identity 74

NCBI Description (U64436) ribosomal protein S8 [Zea mays]

Seq. No. 5243

Contig ID 4620_1.R1040 5'-most EST eep700863857.h1

Method BLASTX
NCBI GI g1710780
BLAST score 663
E value 2.0e-69
Match length 173
% identity 72

NCBI Description 40S RIBOSOMAL PROTEIN S9 (S7) >gi_1321917_emb_CAA65433_

(X96613) cytoplasmic ribosomal protein S7 [Podospora

anserina]

Seq. No. 5244

Contig ID 4620_2.R1040 5'-most EST leu701149795.h1

Method BLASTX
NCBI GI g1710780
BLAST score 669
E value 3.0e-70
Match length 173
% identity 73

NCBI Description 40S RIBOSOMAL PROTEIN S9 (S7) >gi_1321917_emb_CAA65433_

(X96613) cytoplasmic ribosomal protein S7 [Podospora

anserina]

Seq. No. 5245

Contig ID 4620_3.R1040 5'-most EST xpa700793723.h1

Method BLASTX
NCBI GI g1710780
BLAST score 337
E value 2.0e-31
Match length 77
% identity 81

NCBI Description 40S RIBOSOMAL PROTEIN S9 (S7) >gi_1321917_emb_CAA65433_

(X96613) cytoplasmic ribosomal protein S7 [Podospora

anserina]

Seq. No. 5246

Contig ID 4620_4.R1040

5'-most EST LIB3106-058-Q1-K1-H4

Method BLASTX
NCBI GI g131770
BLAST score 132



E value 5.0e-12 Match length 75 % identity 56

NCBI Description 40S RIBOSOMAL PROTEIN S9 (40S RIBOSOMAL PROTEIN 1024) (VEGETATIVE SPECIFIC PROTEIN V12) >gi 70880 pir R3D024

ribosomal protein S9.e - slime mold (Dictyostelium

discoideum) >gi 7353 emb CAA29844 (X06636) rp1024 protein

[Dictyostelium discoideum]

Seq. No. 5247

Contig ID 4625_1.R1040 5'-most EST txt700737518.h1

Method BLASTX
NCBI GI g4049632
BLAST score 759
E value 1.0e-80
Match length 180
% identity 85

NCBI Description (AF039406) pyruvate dehydrogenase kinase [Arabidopsis

thaliana)

Seq. No. 5248

Contig ID 4626 1.R1040

5'-most EST LIB3051-042-Q1-K1-B9

Method BLASTX
NCBI GI g1706551
BLAST score 269
E value 2.0e-23
Match length 115
% identity 57

NCBI Description GLUCAN ENDO-1,3-BETA-GLUCOSIDASE PRECURSOR

((1->3)-BETA-GLUCAN ENDOHYDROLASE) ((1->3)-BETA-GLUCANASE)

(BETA-1,3-ENDOGLUCANASE) >gi_924953 (U30323) beta

1,3-glucanase [Triticum aestivum]

Seq. No. 5249

Contig ID 4626 2.R1040

5'-most EST LIB3092-029-Q1-K1-A12

Method BLASTX
NCBI GI g1706551
BLAST score 926
E value 1.0e-100
Match length 351
% identity 52

NCBI Description GLUCAN ENDO-1,3-BETA-GLUCOSIDASE PRECURSOR

((1->3)-BETA-GLUCAN ENDOHYDROLASE) ((1->3)-BETA-GLUCANASE)

(BETA-1,3-ENDOGLUCANASE) >gi 924953 (U30323) beta

1,3-glucanase [Triticum aestīvum]

Seq. No. 5250

Contig ID 4626 3.R1040

5'-most EST fC-gmro7000749504a1

Method BLASTX
NCBI GI g1706551
BLAST score 230
E value 6.0e-19
Match length 50



% identity 80

NCBI Description GLUCAN ENDO-1, 3-BETA-GLUCOSIDASE PRECURSOR

((1->3)-BETA-GLUCAN ENDOHYDROLASE) ((1->3)-BETA-GLUCANASE)

(BETA-1,3-ENDOGLUCANASE) >gi_924953 (U30323) beta

1,3-glucanase [Triticum aestivum]

Seq. No. 5251

Contig ID 4626 4.R1040

5'-most EST LIB3028-013-Q1-B1-D4

Seq. No. 5252

Contig ID 4630 1.R1040

5'-most EST fC-gmf1700906093j1

Method BLASTX
NCBI GI g4415916
BLAST score 448
E value 4.0e-44
Match length 238
% identity 40

NCBI Description (AC006282) putative pectin methylesterase [Arabidopsis

thaliana]

Seq. No. 5253

Contig ID 4631_1.R1040 5'-most EST seb700653612.h1

Method BLASTX
NCBI GI g3702328
BLAST score 1832
E value 0.0e+00
Match length 443
% identity 72

NCBI Description (AC005397) putative cytochrome b5 [Arabidopsis thaliana]

Seq. No. 5254

Contig ID 4631_2.R1040

5'-most EST jC-gmf102220085g03a1

Method BLASTX
NCBI GI g3819710
BLAST score 462
E value 6.0e-46
Match length 113
% identity 72

NCBI Description (AJ224161) delta-8 sphingolipid desaturase [Arabidopsis

thaliana]

Seq. No. 5255

Contig ID 4638 1.R1040

5'-most EST LIB3028-013-Q1-B1-E8

Method BLASTX
NCBI GI g1931652
BLAST score 712
E value 1.0e-121
Match length 377
% identity 60

NCBI Description (U95973) phosphatidylinositol-4-phosphate 5-kinase isolog

[Arabidopsis thaliana]



Seq. No. 525

Contig ID 4638 2.R1040

5'-most EST jC-gmf102220093g07a1

Method BLASTX
NCBI GI g1931652
BLAST score 336
E value 3.0e-31
Match length 113
% identity 58

NCBI Description (U95973) phosphatidylinositol-4-phosphate 5-kinase isolog

[Arabidopsis thaliana]

Seq. No. 5257

Contig ID 4638 3.R1040

5'-most EST LIB3065-013-Q1-N1-G11

Method BLASTX
NCBI GI g2739367
BLAST score 490
E value 3.0e-49
Match length 142
% identity 63

NCBI Description (AC002505) putative phosphatidylinositol-4-phosphate

5-kinase [Arabidopsis thaliana]

Seq. No. 5258

Contig ID 4641 1.R1040

5'-most EST LIB3028-013-Q1-B1-F10

Seq. No. 5259

Contig ID 4642 1.R1040

5'-most EST LIB3138-045-Q1-N1-D6

Method BLASTX
NCBI GI g3860272
BLAST score 1296
E value 1.0e-143
Match length 270
% identity 90

NCBI Description (AC005824) putative suppressor protein [Arabidopsis

thaliana] >gi 4314399 gb AAD15609 (AC006232) putative skd1

protein [Arabidopsis thaliana]

Seq. No. 5260

Contig ID 4642 2.R1040

5'-most EST uC-gmrominsoy307f06b1

Method BLASTX
NCBI GI g3860272
BLAST score 1219
E value 1.0e-134
Match length 280
% identity 87

NCBI Description (AC005824) putative suppressor protein [Arabidopsis

thaliana] >gi_4314399_gb_AAD15609_ (AC006232) putative skd1

protein [Arabidopsis thaliana]

Seq. No. 5261

Contig ID 4642_3.R1040 5'-most EST g4313953



Method BLASTX
NCBI GI g3860272
BLAST score 1132
E value 1.0e-124
Match length 256
% identity 88

NCBI Description (AC005824) putative suppressor protein [Arabidopsis

thaliana] >gi 4314399_gb_AAD15609_ (AC006232) putative skdl

protein [Arabidopsis Thaliana]

Seq. No. 5262

Contig ID 4642_4.R1040
5'-most EST g4314039
Method BLASTX
NCBI GI g3860272
BLAST score 478
E value 9.0e-48
Match length 103
% identity 83

% identity 83
NCBI Description (AC005824) putative suppressor protein [Arabidopsis

thaliana] >gi_4314399_gb_AAD15609_ (AC006232) putative skd1

protein [Arabidopsis thaliana]

Seq. No. 5263

Contig ID 4642_9.R1040 5'-most EST pmv700893510.h1

Method BLASTX
NCBI GI g3860272
BLAST score 203
E value 4.0e-16
Match length 82
% identity 68

NCBI Description (AC005824) putative suppressor protein [Arabidopsis

thaliana] >gi_4314399_gb_AAD15609_ (AC006232) putative skd1

protein [Arabidopsis thaliana]

Seq. No. 5264

Contig ID 4645_1.R1040

5'-most EST LIB3170-057-Q1-K1-C8

Method BLASTX
NCBI GI g3293547
BLAST score 174
E value 3.0e-12
Match length 141
% identity 34

NCBI Description (AF072709) putative oxidoreductase [Streptomyces lividans]

Seq. No. 5265

Contig ID 4645_2.R1040

5'-most EST LIB3028-013-Q1-B1-F4

Method BLASTX
NCBI GI g117549
BLAST score 188
E value 6.0e-14
Match length 173
% identity 34

NCBI Description QUINONE OXIDOREDUCTASE (NADPH:QUINONE REDUCTASE)



(ZETA-CRYSTALLIN) >gi 65895 pir__CYGPZ zeta-crystallin / quinone reductase (NADPH) (EC 1.6.-.-) - guinea pig >gi 305333 (M26936) zeta-crystallin [Cavia porcellus]

5266 Seq. No. Contig ID 4646 1.R1040 kl1701208545.h1 5'-most EST BLASTX Method g4079809 NCBI GI

350 BLAST score 7.0e-33 E value 190 Match length 3 % identity

NCBI Description (AF071172) HERC2 [Homo sapiens]

5267 Seq. No.

4649 1.R1040 Contig ID asn701137915.h1 5'-most EST

5268 Seq. No.

4649 2.R1040 Contig ID

LIB3050-001-Q1-E1-F7 5'-most EST

BLASTX Method NCBI GI g416650 509 BLAST score 3.0e-51E value 210 Match length 48 % identity

PROBABLE GLUTATHIONE S-TRANSFERASE (AUXIN-INDUCED PROTEIN NCBI Description

PGNT35/PCNT111) >gi 100304 pir S16268 auxin-induced

protein (clone pGNT35) - common tobacco

>qi 19797 emb CAA39706 (X56265) auxin-induced protein [Nicotiana tabacum] >gi 19801_emb CAA39710_ (X56269)

auxin-induced protein [Nicotiana tabacum]

5269 Seq. No.

4651 1.R1040 Contig ID

5'-most EST LIB3028-013-Q1-B1-C7

5270 Seq. No.

4652 1.R1040 Contig ID

5'-most EST LIB3139-052-P1-N1-G3

Method BLASTX NCBI GI q3204134 BLAST score 2063 E value 0.0e + 00Match length 677 % identity 90

(AJ006771) beta-galactosidase [Cicer arietinum] NCBI Description

Seq. No.

5271

4655 1.R1040 Contig ID

5'-most EST uC-gmronoir039g01b1

Seq. No.

5272

4659 1.R1040 Contig ID

5'-most EST

LIB3028-013-Q1-B1-D7



```
BLASTX
Method
                  g4107099
NCBI GI
                  247
BLAST score
                  2.0e-20
E value
                  149
Match length
% identity
                  (AB015141) AHP1 [Arabidopsis thaliana]
NCBI Description
                  >gi 4156245 dbj BAA37112 (AB012570) ATHP3 [Arabidopsis
                   thaliana]
                   5273
Seq. No.
                   4662 1.R1040
Contig ID
                   wvk700681920.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q951449
BLAST score
                   1796
E value
                   0.0e + 00
                   507
Match length
% identity
                   67
                  (L46681) aspartic protease precursor [Lycopersicon
NCBI Description
                   esculentum]
                   5274
Seq. No.
                   4662 2.R1040
Contig ID
                   LIB3028-009-Q1-B1-B10
5'-most EST
                   BLASTX
Method
NCBI GI
                   q951449
                   933
BLAST score
                   1.0e-101
E value
Match length
                   249
% identity
                   (L46681) aspartic protease precursor [Lycopersicon
NCBI Description
                   esculentum]
                   5275
Seq. No.
                   4663 1.R1040
Contig ID
                   LIB3093-051-Q1-K1-A12
5'-most EST
                   BLASTX
Method
                   g4454467
NCBI GI
                   414
BLAST score
                   3.0e-40
E value
                   84
Match length
 % identity
                   85
                   (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5276
 Seq. No.
```

4664 1.R1040 Contig ID leu701154408.h1 5'-most EST

BLASTX Method q2244732 NCBI GI BLAST score 840 3.0e-90 E value 182 Match length % identity

(D88413) endo-xyloglucan transferase [Gossypium hirsutum] NCBI Description

5277 Seq. No.

```
4668 1.R1040
Contig ID
5'-most EST
                   uC-gmronoir013a07b1
Seq. No.
                   4668 2.R1040
Contig ID
5'-most EST
                   qsv701056644.hl
Seq. No.
                   5279
Contig ID
                   4672 1.R1040
5'-most EST
                   leu701153884.h1
Method
                   BLASTX
NCBI GI
                   q4105097
BLAST score
                   342
E value
                   2.0e-32
                   79
Match length
                   82
% identity
                  (AF043255) MADS box protein 26 [Cucumis sativus]
NCBI Description
Seq. No.
                   5280
Contig ID
                   4674 1.R1040
                   leu701149025.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2245109
BLAST score
                   589
                   5.0e-61
E value
Match length
                   145
% identity
                   77
NCBI Description
                   (Z97343) carboxyl-terminal proteinase homolog [Arabidopsis
                   thaliana]
                   5281
Seq. No.
                   4675 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy083c10b1
Method
                   BLASTX
NCBI GI
                   g3219271
BLAST score
                   886
E value
                   2.0e-95
Match length
                   224
                   76
% identity
NCBI Description
                   (AB015315) MAP kinase kinase 4 [Arabidopsis thaliana]
                   5282
Seq. No.
                   4675 2.R1040
Contig ID
5'-most EST
                   LIB3051-015-Q1-E1-F4
Method
                   BLASTX
NCBI GI
                   g3219271
BLAST score
                   762
E value
                   6.0e-81
                   222
Match length
```

71 % identity

(AB015315) MAP kinase kinase 4 [Arabidopsis thaliana] NCBI Description

Seq. No.

Contig ID 4675 3.R1040

5'-most EST jC-gmro02800039d05a1

5283

Method BLASTX NCBI GI q3219271



185 BLAST score 3.0e-15 E value 97 Match length 60 % identity

(AB015315) MAP kinase kinase 4 [Arabidopsis thaliana] NCBI Description

5284 Seq. No.

4675 4.R1040 Contig ID

jC-gmf102220052d10d1 5'-most EST

Method BLASTX q1523800 NCBI GI 162 BLAST score 4.0e-11 E value 51 Match length 57 % identity

(Y07694) MAP kinase kinase alpha protein kinase NCBI Description

[Arabidopsis thaliana]

Seq. No. 5285

4676 1.R1040 Contig ID cf1700863656.h1 5'-most EST

BLASTX Method q2392895 NCBI GI 1512 BLAST score 1.0e-169 E value 333 Match length % identity

(AF017056) brassinosteroid insensitive 1 [Arabidopsis NCBI Description

thaliana]

5286 Seq. No.

4677 1.R1040 Contig ID ssr700557419.h15'-most EST

BLASTX Method q2072986 NCBI GI 183 BLAST score 2.0e-13 E value 67 Match length % identity 57

(U95142) putative G-protein-coupled receptor [Arabidopsis NCBI Description

thaliana] >gi_2072988 (U95143) putative G-protein-coupled

receptor [Arabidopsis thaliana]

5287 Seq. No.

4677 2.R104€ Contig ID

LIB3109-034-Q1-K1-F6 5'-most EST

BLASTX Method g2072986 NCBI GI BLAST score 462 3.0e-46 E value 118 Match length 74 % identity

(U95142) putative G-protein-coupled receptor [Arabidopsis NCBI Description

thaliana] >gi_2072988 (U95143) putative G-protein-coupled

receptor [Arabidopsis thaliana]

5288 Seq. No.



Contig ID 4679_1.R1040 5'-most EST leu701148521.h1

Method BLASTX
NCBI GI g4099833
BLAST score 559
E value 2.0e-57
Match length 186
% identity 56

NCBI Description (U90265) bifunctional nuclease [Zinnia elegans]

Seq. No. 5289

Contig ID 4680_1.R1040 5'-most EST fC-gmse700674289f1

Seq. No. 5290

Contig ID 4680_2.R1040

5'-most EST uC-gmrominsoy227f07b1

Seq. No. 5291

Contig ID 4680_4.R1040 5'-most EST zsg701117722.h2

Seq. No. 5292

Contig ID 4682_1.R1040

5'-most EST LIB3028-012-Q1-B1-H1

Seq. No. 5293

Contig ID 4696_1.R1040 5'-most EST hyd700727117.h1 Method BLASTN

Method BLASTN
NCBI GI g4519194
BLAST score 63
E value 1.0e-26
Match length 227

Match length 227 % identity 82

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MHM17, complete sequence

Seq. No. 5294

Contig ID 4696 2.R1040

5'-most EST LIB3051-004-Q1-E1-E3

Method BLASTN
NCBI GI g4519194
BLAST score 48
E value 1.0e-17
Match length 176
% identity 82

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MHM17, complete sequence

Seq. No. 5295

Contig ID 4697_1.R1040

5'-most EST LIB3028-010-Q1-B1-F2

Seq. No. 5296

Contig ID 4700 1.R1040

5'-most EST LIB3074-019-Q1-E1-C2

Match length

112



```
BLASTX
Method
                  g4325369
NCBI GI
                  251
BLAST score
                  2.0e-21
E value
                  81
Match length
% identity
NCBI Description (AF128396) T3H13.3 gene product [Arabidopsis thaliana]
                  5297
Seq. No.
                  4701 1.R1040
Contig ID
                  q439\overline{5}911
5'-most EST
                  BLASTX
Method
                  g730542
NCBI GI
BLAST score
                  245
                   2.0e-20
E value
                   98
Match length
                   54
% identity
                   60S RIBOSOMAL PROTEIN L22 (HEPARIN BINDING PROTEIN HBP15)
NCBI Description
                   >gi_543098_pir__JC2121 heparin-binding protein 15 - pig
                   >gi_627874_pir__JC2119 heparin-binding protein 15 - mouse
                   >gi_409072_dbj_BAA04546_ (D17653) HBp15/L22 [Mus musculus]
                   >gi 409074 dbj BAA04547 (D17654) HBp15/L22 [Sus scrofa]
                   5298
Seq. No.
                   4701 2.R1040
Contig ID
                   LIB3073-025-Q1-K1-G2
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1172995
                   249
BLAST score
                   4.0e-21
E value
Match length
                   101
                   52
% identity
                   60S RIBOSOMAL PROTEIN L22 >gi_1083790_pir__S52084 ribosomal
NCBI Description
                   protein L22 - rat >gi 710295 emb CAA55204 (X78444)
                   ribosomal protein L22 [Rattus norvegicus]
                   >gi 1093952_prf _2105193A ribosomal protein L22 [Rattus
                   norvegicus]
                   5299
Seq. No.
                   4703 1.R1040
Contig ID
5'-most EST
                   uaw700666912.h1
Method
                   BLASTX
NCBI GI
                   g2522534
BLAST score
                   519
                   1.0e-105
E value
                   415
Match length
% identity
                   29
                   (AF027302) TNF-alpha stimulated ABC protein [Homo sapiens]
NCBI Description
                   5300
Seq. No.
                   4704 1.R1040
Contig ID
5'-most EST
                   LIB3106-037-Q1-K1-D6
Method
                   BLASTX
                   q4371280
NCBI GI
BLAST score
                   426
                   9.0e-42
E value
```

BLAST score

Match length

NCBI Description

% identity

E value

386

83

87

3.0e-37



```
% identity
                   76
NCBI Description
                   (AC006260) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   5301
                   4710 2.R1040
Contig ID
5'-most EST
                  LIB3028-012-Q1-B1-E7
Method
                  BLASTX
NCBI GI
                   q3877951
BLAST score
                   246
E value
                   1.0e-20
Match length
                   133
% identity
                   36
                   (Z81555) predicted using Genefinder [Caenorhabditis
NCBI Description
                   elegans]
Seq. No.
                   5302
Contig ID
                   4712 1.R1040
5'-most EST
                  LIB3051-017-Q1-E1-A4
Seq. No.
                   5303
Contig ID
                   4714 1.R1040
5'-most EST
                  LIB3109-035-Q1-K6-H10
Method
                  BLASTX
NCBI GI
                   g3892058
BLAST score
                   490
E value
                   9.0e-49
Match length
                   226
% identity
                   47
NCBI Description
                   (AC002330) putative glutamate-/aspartate-binding peptide
                   [Arabidopsis thaliana]
                   5304
Seq. No.
                   4714 2.R1040
Contig ID
5'-most EST
                  LIB3139-030-P1-N1-B7
Method
                  BLASTX
NCBI GI
                   g2244904
BLAST score.
                   181
E value
                   3.0e-13
Match length
                   85
% identity
                   62
NCBI Description
                   (Z97339) similar to hypothetical protein C02F5.7 - Caenorha
                   [Arabidopsis thaliana]
Seq. No.
                   5305
                   4714 3.R1040
Contig ID
5'-most EST
                  uC-gmronoir030b08b1
Seq. No.
                   5306
                   4716 1.R1040
Contiq ID
5'-most EST
                  leu701154646.h1
                  BLASTX
Method
NCBI GI
                  q4262183
```

1016

(AC005508) 51434 [Arabidopsis thaliana]



```
Seq. No.
                  5307
Contig ID
                  4718 1.R1040
5'-most EST
                  LIB3028-012-Q1-B1-D5
Method
                  BLASTX
NCBI GI
                  q4220480
BLAST score
                  242
E value
                  3.0e-20
Match length
                  101
% identity
                  51
                  (AC006069) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  5308
                  4721 1.R1040
Contig ID
                  txt700734653.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4454051
BLAST score
                  581
                  6.0e-60
E value
                  208
Match length
% identity
NCBI Description
                  (AL035394) putative polygalacturonase [Arabidopsis
                  thaliana]
Seq. No.
                  5309
Contig ID
                  4724 1.R1040
5'-most EST
                  jC-gmf102220115f09a1
Method
                  BLASTX
NCBI GI
                  g2739370
BLAST score
                  333
E value
                  4.0e-31
Match length
                  128
% identity
                  51
NCBI Description
                  (AC002505) putative pectinesterase [Arabidopsis thaliana]
Seq. No.
                  5310
                  4725 1.R1040
Contig ID
                  LIB3028-007-Q1-B1-G3
5'-most EST
Method
                  BLASTX
NCBI GI
                  g602292
                  507
BLAST score
E value
                  3.0e-51
                  208
Match length
                  49
% identity
NCBI Description
                  (U17987) RCH2 protein [Brassica napus]
Seq. No.
                  5311
Contig ID
                  4729 1.R1040
5'-most EST
                  crh700852073.h1
Method
                  BLASTX
```

Method BLASTX
NCBI GI 94406759
BLAST score 542
E value 4.0e-55
Match length 291
% identity 44

NCBI Description (AC006836) hypothetical protein [Arabidopsis thaliana]



5312 Seq. No.

Contig ID 4729 2.R1040 5'-most EST kl1701212931.h1

BLASTX Method q4406759 NCBI GI BLAST score 362 6.0e - 34E value 239 Match length % identity 36

(AC006836) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 5313

Contig ID 4733 1.R1040

5'-most EST uC-gmropic019c04b1

Method BLASTX NCBI GI g1350983 BLAST score 1092 E value 1.0e-119 Match length 234 % identity 88

NCBI Description 40S RIBOSOMAL PROTEIN S3A (CYC07 PROTEIN)

Seq. No. 5314

Contig ID 4737 1.R1040

5'-most EST LIB3028-012-Q1-B1-B12

Seq. No. 5315

Contig ID 4739 1.R1040

5'-most EST LIB3170-025-Q1-K1-H1

Method BLASTX NCBI GI q3892058 BLAST score 832 E value 4.0e-89 Match length 238 % identity 66

NCBI Description

(AC002330) putative glutamate-/aspartate-binding peptide

[Arabidopsis thaliana]

Seq. No. 5316

Contig ID 4740 1.R1040 5'-most EST zzp700834267.h1

Method BLASTX NCBI GI g2392021 BLAST score 628 E value 3.0e-65 Match length 168 % identity 68

(D63425) phopholipid hydroperoxide glutathione NCBI Description

peroxidase-like protein [Spinacia oleracea]

Seq. No. 5317

Contig ID 4743 1.R1040

5'-most EST LIB3109-019-Q1-K1-D8

Method BLASTX NCBI GI q2460203 BLAST score 268 E value 1.0e-23



```
Match length
                   47
% identity
                   (AF021244) coronatine-induced protein 1 [Arabidopsis
NCBI Description
                   thaliana]
                   5318
Seq. No.
                   4744 1.R1040
Contig ID
                   zhf700956861.h1
5'-most EST
                   BLASTX
Method
                   q3193286
NCBI GI
BLAST score
                   199
                   3.0e-15
E value
                   54
Match length
                   63
% identity
NCBI Description (AF069298) T14P8.22 gene product [Arabidopsis thaliana]
                   5319
Seq. No.
                   4744 2.R1040
Contig ID
5'-most EST
                   LIB3074-012-Q1-E1-E10
                   5320
Seq. No.
                   4744 3.R1040
Contig ID
                   LIB3138-007-Q1-N1-D10
5'-most EST
                   BLASTX
Method
                   g3193286
NCBI GI
                   587
BLAST score
                   2.0e-60
E value
                   255
Match length
% identity
NCBI Description (AF069298) T14P8.22 gene product [Arabidopsis thaliana]
Seq. No.
                   4745 1.R1040
Contig ID
                   uC-gmflminsoy011h05b1
5'-most EST
                   5322
Seq. No.
                   4746 1.R1040
Contig ID
                   uC-gmrominsoy040e05b1
5'-most EST
                   5323
Seq. No.
                   4746 2.R1040
Contig ID
5'-most EST
                   kl1701202777.h1
                   5324
Seq. No.
                    4747 1.R1040
Contig ID
                    jsh7\overline{0}1064513.h1
5'-most EST
                   BLASTX
Method
                    g2088643
NCBI GI
                    509
BLAST score
                    1.0e-84
E value
```

Match length % identity

45 (AF002109) transcription factor SF3 isolog [Arabidopsis NCBI Description

thaliana]

197

5325

Seq. No.

4747 2.R1040 Contig ID

1019

Contiq ID

5'-most EST

```
5'-most EST
                  LIB3139-076-P1-N1-C2
Method
                  BLASTX
NCBI GI
                  q1841464
BLAST score
                   334
E value
                   5.0e-31
Match length
                  75
% identity
                   42
NCBI Description
                  (Y11002) LIM-domain SF3 protein [Nicotiana tabacum]
Seq. No.
                   5326
                   4747 4.R1040
Contig ID
5'-most EST
                  sat701015226.h1
                  5327
Seq. No.
                  4747 5.R1040
Contig ID
5'-most EST
                  pmv700892040.hl
                  5328
Seq. No.
Contig ID
                  4749 1.R1040
5'-most EST
                  crh700852401.h1
                  BLASTX
Method
NCBI GI
                  q480618
BLAST score
                  724
E value
                  2.0e-76
Match length
                  216
% identity
                   69
NCBI Description
                  ATAF1 protein - Arabidopsis thaliana (fragment)
                  >gi_1345506_emb_CAA52771_ (X74755) ATAF1 [Arabidopsis
                  thaliana]
Seq. No.
                   5329
Contig ID
                   4749 2.R1040
5'-most EST
                  uC-gmrominsoy245e10b1
Method
                  BLASTX
                   g480618
NCBI GI
BLAST score
                   677
E value
                   6.0e-71
Match length
                   206
% identity
                   68
                  ATAF1 protein - Arabidopsis thaliana (fragment)
NCBI Description
                   >gi_1345506_emb_CAA52771_ (X74755) ATAF1 [Arabidopsis
                   thaliana]
Seq. No.
                   5330
                   4749 3.R1040
Contig ID
5'-most EST
                   uC-gmronoir076d10b1
Method
                  BLASTX
                   q4218535
NCBI GI
BLAST score
                   446
                  2.0e-44
E value
                  126
Match length
                   64
% identity
NCBI Description (AJ010829) GRAB1 protein [Triticum sp.]
Seq. No.
                   5331
```

4753 1.R1040

LIB3109-046-Q1-K1-D11



Method BLASTX
NCBI GI g4455363
BLAST score 471
E value 5.0e-47
Match length 216
% identity 46

NCBI Description (AL035524) Medicago nodulin N21-like protein [Arabidopsis

thaliana]

Seq. No. 5332

Contig ID 4754 1.R1040

5'-most EST LIB3039-009-Q1-E1-B7

Seq. No. 5333

Contig ID 4758 1.R1040

5'-most EST LIB3072-056-Q1-K1-E2

Method BLASTN
NCBI GI g3687405
BLAST score 72
E value 4.0e-32
Match length 148
% identity 87

NCBI Description Lycopersicon esculentum mRNA for hypothetical protein

Seq. No. 5334

Contig ID 4758 2.R1040 g4290065 5'-most EST Method BLASTN NCBI GI g3687405 BLAST score 82 E value 4.0e-38 Match length 146 % identity 89

NCBI Description Lycopersicon esculentum mRNA for hypothetical protein

Seq. No. 5335

Contig ID 4758_3.R1040 5'-most EST jsh701063726.h1

Method BLASTN
NCBI GI g3687405
BLAST score 79
E value 2.0e-36
Match length 139
% identity 89

NCBI Description Lycopersicon esculentum mRNA for hypothetical protein

Seq. No. 5336

Contig ID 4758_4.R1040 5'-most EST zhf700962928.h1

Method BLASTN
NCBI GI g3297806
BLAST score 34
E value 1.0e-09
Match length 66
% identity 47

NCBI Description Arabidopsis thaliana DNA chromosome 4, BAC clone F1715

(ESSAII project)



5337 Seq. No.

4763 1.R1040 Contig ID $k117\overline{0}1214127.h1$ 5'-most EST

5338 Seq. No.

4763 2.R1040 Contig ID 5'-most EST pxt700945468.hl

Seq. No. 5339

4766 1.R1040 Contig ID

uC-gmronoir042a03b1 5'-most EST

BLASTX Method q4406809 NCBI GI BLAST score 245 3.0e-20 E value Match length 295 31 % identity

NCBI Description (AC006201) unknown protein [Arabidopsis thaliana]

5340 Seq. No.

4766 2.R1040 Contig ID 5'-most EST awf700837035.hl

Method BLASTN q4097879 NCBI GI BLAST score 35 2.0e-10 E value Match length 51 92 % identity

Bean pod mottle virus complete segment RNA1 polyprotein NCBI Description

gene, complete cds

5341 Seq. No.

4770 1.R1040 Contig ID

jC-gmle01810064h08a1 5'-most EST

BLASTX Method g1086225 NCBI GI 949 BLAST score 0.0e + 00E value 534 Match length 76 % identity

RING-finger protein (C-terminal) - Lotus japonicus NCBI Description

>gi_558545_emb_CAA85321_ (Z36750) protein containing

C-terminal RING-finger [Lotus japonicus]

>gi_1771195_emb_CAA70734_ (Y09539) RING-finger protein

[Lotus japonicus]

5342 Seq. No.

4775 1.R1040 Contig ID 5'-most EST pxt700943412.h1

Seq. No. 5343

4777 1.R1040 Contig ID

5'-most EST LIB3109-021-Q1-K2-E3

Method BLASTX q1531758 NCBI GI BLAST score 750



E value 5.0e-80 Match length 147 % identity 92

(X98772) AUX1 [Arabidopsis thaliana] >gi 3335360 (AC003028) NCBI Description

unknown protein [Arabidopsis thaliana]

Seq. No.

4782 1.R1040 Contig ID

5'-most EST LIB3028-011-Q1-B1-E6

5344

Seq. No.

5345 4782 2.R1040 Contig ID

5'-most EST LIB3055-005-Q1-N1-F10

5346

Seq. No.

Contig ID 4783 1.R1040 leu701157320.h1 5'-most EST

Seq. No.

5347

4783 2.R1040 Contig ID

5'-most EST jC-gmle01810011d10a1

Seq. No.

5348

4786 1.R1040 Contig ID

5'-most EST LIB3139-064-P1-N1-H5

Seq. No. 5349

4787 1.R1040 Contig ID

5'-most EST LIB3028-011-Q1-B1-B12

Method BLASTX NCBI GI q1914845 BLAST score 383 1.0e-36 E value 135 Match length

% identity 59

NCBI Description (U89496) liguleless1 protein [Zea mays]

Seq. No. 5350

4788_1.R1040 Contig ID

5'-most EST jC-gmle01810042c06a1

Method BLASTX NCBI GI g2244971 BLAST score 365 E value 1.0e-34 131 Match length % identity 63

NCBI Description (297340) hypothetical protein [Arabidopsis thaliana]

Seq. No.

5351

4788 2.R1040 Contig ID

5'-most EST uC-gmflminsoy058c06b1

Method BLASTX NCBI GI g2244971 BLAST score 508 1.0e-51 E value Match length 154 % identity 64



```
(Z97340) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  5352
Seq. No.
Contig ID
                  4791 1.R1040
5'-most EST
                  LIB3138-030-Q1-N1-H11
Seq. No.
                  5353
Contig ID
                  4792 1.R1040
5'-most EST
                  jC-gmf102220050e08a1
Method
                  BLASTX
NCBI GI
                  q2827552
BLAST score
                  804
E value
                  3.0e-86
Match length
                  173
% identity
                  41
NCBI Description (AL021635) predicted protein [Arabidopsis thaliana]
Seq. No.
                  5354
Contig ID
                  4799 1.R1040
5'-most EST
                  LIB3073-020-Q1-K1-C6
Method
                  BLASTX
NCBI GI
                  g4406772
BLAST score
                  156
E value
                  4.0e-10
Match length
                  63
% identity
                  60
NCBI Description
                  (AC006836) putative nitrilase-associated protein
                   [Arabidopsis thaliana]
Seq. No.
                  5355
Contiq ID
                  4799 2.R1040
5'-most EST
                  LIB3049-003-Q1-E1-E9
Seq. No.
                  5356
Contig ID
                  4801 1.R1040
5'-most EST
                  fde700873950.h1
Seq. No.
                  5357
Contig ID
                  4802_1.R1040
5'-most EST
                  LIB3107-080-Q1-K1-C1
Method
                  BLASTN
NCBI GI
                  q1813328
BLAST score
                  212
E value
                  1.0e-115
Match length
                  671
% identity
NCBI Description Canavalia gladiata mRNA for HMG-1, complete cds
Seq. No.
                  5358
```

4802 2.R1040 Contig ID 5'-most EST pcp700991734.hl

Method BLASTN NCBI GI g1813328 BLAST score 126 E value 2.0e-64 Match length 286 % identity 86



NCBI Description Canavalia gladiata mRNA for HMG-1, complete cds

Seq. No.

4803 1.R1040 Contig ID

5'-most EST LIB3028-010-Q1-B1-H6

5359

Method BLASTX NCBI GI g100347 BLAST score 498 E value 1.0e-77 Match length 220 66 % identity

NCBI Description monosaccharide transport protein MST1 - common tobacco

>gi_19885_emb_CAA47324 (X66856) monosaccharid transporter

[Nicotiana tabacum]

Seq. No. 5360

4803 2.R1040 Contig ID

5'-most EST jC-gmf102220138h04d1

Method BLASTX NCBI GI g99758 BLAST score 344 2.0e-32 E value Match length 92 % identity 65

NCBI Description monosaccharid transport protein STP4 - Arabidopsis thaliana

>gi 16524 emb CAA47325 (X66857) sugar transport protein

[Arabidopsis thaliana]

Seq. No.

5361 Contig ID

4805 1.R1040

5'-most EST uC-gmrominsoy053f06b1

Method BLASTX NCBI GI g3914359 BLAST score 1743 E value 0.0e + 00Match length 362 90 % identity

PHOSPHOLIPASE D PRECURSOR (PLD) (CHOLINE PHOSPHATASE) NCBI Description

(PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D)

>gi 1928979 (U92656) phospholipase D [Vigna unguiculata]

5362 Seq. No.

4805 2.R1040 Contig ID 5'-most EST leu701151758.h1

Method BLASTN NCBI GI g1928978 BLAST score 181 E value 3.0e-97 Match length 282 93 % identity

Vigna unquiculata phospholipase D mRNA, complete cds NCBI Description

Seq. No. 5363

Contig ID 4806 1.R1040 5'-most EST zsg701118351.h1

Method BLASTX NCBI GI q4220521



BLAST score 780 E value 8.0e-83 Match length 345 % identity 57

NCBI Description (AL035356) putative protein [Arabidopsis thaliana]

Seq. No. 5364

Contig ID 4806_4.R1040 5'-most EST zsg701118740.h1

Seq. No. 5365

Contig ID 4809_1.R1040

5'-most EST LIB3028-011-Q1-B1-A5

Method BLASTX
NCBI GI g1769895
BLAST score 803
E value 6.0e-86
Match length 228
% identity 68

NCBI Description (X96598) CaLB protein [Arabidopsis thaliana]

Seq. No. 5366

Contig ID 4811_1.R1040

5'-most EST uC-gmrominsoy114g12b1

Method BLASTX
NCBI GI g1361979
BLAST score 310
E value 5.0e-30
Match length 108
% identity 70

NCBI Description serine O-acetyltransferase (EC 2.3.1.30) - watermelon

>gi_1350550_dbj_BAA12843_ (D85624) serine acetyltransferase

[Citrullus lanatus] >gi_1841312_dbj_BAA08479_ (D49535)

serine acetyltransferase. [Citrullus lanatus] >gi_2337772_dbj_BAA21827_ (AB006530) serine

acetyltransferase [Citrullus lanatus]

Seq. No. 5367

Contig ID 4814_1.R1040 5'-most EST zhf700954244.h1

Method BLASTX
NCBI GI g3377820
BLAST score 265
E value 8.0e-37
Match length 132
% identity 62

NCBI Description (AF076275) contains similarity to coatomer zeta chains

[Arabidopsis thaliana]

Seq. No. 5368

Contig ID 4815_1.R1040

5'-most EST LIB3109-034-Q1-K1-E7

Method BLASTX
NCBI GI g2454184
BLAST score 1235
E value 1.0e-136
Match length 405

1026



```
% identity
                   (U80186) pyruvate dehydrogenase E1 beta subunit
NCBI Description
                   [Arabidopsis thaliana]
                  5369
Seq. No.
                   4815 2.R1040
Contig ID
5'-most EST
                  seb700650890.h1
                   5370
Seq. No.
                   4815 3.R1040
Contig ID
                   jC-gmro02910071g01d1
```

5371 Seq. No.

5'-most EST

4816 1.R1040 Contig ID LIB3028-010-Q1-B1-F9

5'-most EST BLASTX Method g2760324 NCBI GI 322 BLAST score 1.0e-29 E value

104 Match length 64 % identity

(AC002130) F1N21.9 [Arabidopsis thaliana] NCBI Description

Seq. No.

4817 1.R1040 Contig ID

LIB3040-054-Q1-E1-B9 5'-most EST

5373 Seq. No.

4817 2.R1040 Contig ID

LIB3072-027-Q1-E1-D9 5'-most EST

5374 Seq. No.

4820 1.R1040 Contig ID $leu7\overline{0}1147178.h1$ 5'-most EST

BLASTX Method q4218120 NCBI GI 1243 BLAST score 1.0e-137 E value 356 Match length % identity

(AL035353) Proline-rich APG-like protein [Arabidopsis NCBI Description

thaliana]

Seq. No. 5375

4820 2.R1040 Contig ID wrg700788041.h1 5'-most EST

BLASTX Method g4218120 NCBI GI BLAST score 498 3.0e-50 E value 122 Match length 77 % identity

(AL035353) Proline-rich APG-like protein [Arabidopsis NCBI Description

thaliana]

5376 Seq. No.

4820_3.R1040 Contig ID

1027



```
q4260376
5'-most EST
                  BLASTX
Method
                   g4218120
NCBI GI
                   331
BLAST score
                   8.0e-31
E value
                   79
Match length
% identity
                   (AL035353) Proline-rich APG-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   5377
Seq. No.
                   4823 1.R1040
Contig ID
                   jC-qmro02800040g03d1
5'-most EST
                   5378
Seq. No.
                   4823 2.R1040
Contig ID
5'-most EST
                   uxk7\overline{0}0672939.h1
                   BLASTX
Method
                   g3650033
NCBI GI
                   200
BLAST score
                   2.0e-15
E value
Match length
                   81
                   54
% identity
                  (AC005396) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5379
Seq. No.
                   4823 3.R1040
Contig ID
                   jC-gmf102220143c10d1
5'-most EST
                   5380
Seq. No.
                   4826 1.R1040
Contig ID
                   LIB3093-007-Q1-K1-E3
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3415115
BLAST score
                   498
                   7.0e-50
E value
                   253
Match length
                   48
% identity
                   (AF081202) villin 2 [Arabidopsis thaliana]
NCBI Description
                   5381
Seq. No.
                   4827 1.R1040
Contig ID
5'-most EST
                   LIB3028-010-Q1-B1-E6
Method
                   BLASTX
NCBI GI
                   q3618316
                   188
BLAST score
                    4.0e-14
E value
                   100
Match length
% identity
                   (AB001886) zinc finger protein [Oryza sativa]
NCBI Description
                    5382
Seq. No.
Contig ID
                    4828 1.R1040
                    wvk700680503.hl
5'-most EST
```

1028

BLASTX

236

q4539010

Method

NCBI GI

BLAST score



E value 1.0e-19
Match length 143
% identity 48

NCBI Description (AL049481) putative DNA-directed RNA polymerase

[Arabidopsis thaliana]

Seq. No. 5383

Contig ID 4830_1.R1040

5'-most EST LIB3106-035-Q1-K1-B7

Method BLASTX
NCBI GI g4138916
BLAST score 979
E value 1.0e-106
Match length 242
% identity 74

NCBI Description (AF059489) expansin precursor [Lycopersicon esculentum]

Seq. No. 5384

Contig ID 4830_2.R1040 5'-most EST hyd700724977.h1

Method BLASTX
NCBI GI 94027897
BLAST score 452
E value 6.0e-45
Match length 116
% identity 71

NCBI Description (AF049353) alpha-expansin precursor [Nicotiana tabacum]

Seq. No. 5385

Contig ID 4834_1.R1040

5'-most EST LIB3138-039-Q1-N1-C11

Method BLASTX
NCBI GI 94468813
BLAST score 748
E value 3.0e-79
Match length 200
% identity 67

NCBI Description (AL035601) putative protein [Arabidopsis thaliana]

Seq. No. 5386

Contig ID 4836 1.R1040

5'-most EST jC-gmro02910031e04a1

Seq. No. 5387

Contig ID 4837 1.R1040

5'-most EST LIB3049-053-Q1-E1-E8

Seq. No. 5388

Contig ID 4837_2.R1040 5'-most EST trc700566031.h1

Seq. No. 5389

Contig ID 4840_1.R1040

5'-most EST LIB3028-010-Q1-B1-D10

Seq. No. 5390

Contig ID 4844 1.R1040

BLAST score

428



```
LIB3028-010-Q1-B1-B9
5'-most EST
                  5391
Seq. No.
                  4845 1.R1040
Contig ID
                  LIB3138-126-Q1-N1-A8
5'-most EST
Seq. No.
                  5392
                   4846 1.R1040
Contig ID
                  uC-gmropic014h09b1
5'-most EST
                  BLASTX
Method
                  q1546702
NCBI GI
                  819
BLAST score
                   1.0e-87
E value
                   248
Match length
                   62
% identity
NCBI Description (X98809) peroxidase ATP5a [Arabidopsis thaliana]
                   5393
Seq. No.
                   4846 2.R1040
Contig ID
                   bth700844765.h1
5'-most EST
                   BLASTX
Method
                   g1546702
NCBI GI
                   542
BLAST score
                   2.0e-55
E value
                   151
Match length
                   68
% identity
NCBI Description (X98809) peroxidase ATP5a [Arabidopsis thaliana]
                   5394
Seq. No.
                   4846 3.R1040
Contig ID
                   pmv700893877.h1
5'-most EST
                   BLASTX
Method
                   g2811262
NCBI GI
                   121
BLAST score
                   2.0e-11
E value
                   75
Match length
                   56
% identity
                   (AF043234) ferriprotein porphyrin-containing peroxidase
NCBI Description
                   [Striga asiatica]
                   5395
Seq. No.
                   4848 1.R1040
Contig ID
5'-most EST
                   LIB3028-010-Q1-B1-A9
Method
                   BLASTN
                   g2746720
NCBI GI
BLAST score
                   128
                   2.0e-65
E value
Match length
                   312
% identity
                   85
                   Capsicum annuum histone H4 mRNA, complete cds
NCBI Description
                   5396
Seq. No.
                   4850 1.R1040
 Contig ID
                   g5687831
 5'-most EST
                   BLASTN
Method
                   g3127889
NCBI GI
```



E value 0.0e+00
Match length 760
% identity 89

NCBI Description Cicer arietinum mRNA for cysteine synthase, partial

Seq. No. 5397

Contig ID 4850_2.R1040

5'-most EST LIB3087-002-Q1-K1-A3

Method BLASTN
NCBI GI g3127889
BLAST score 149
E value 4.0e-78
Match length 193
% identity 94

NCBI Description Cicer arietinum mRNA for cysteine synthase, partial

Seq. No.

5398

Contig ID

4856 1.R1040

5'-most EST

LIB3106-003-Q1-K1-F1

Method BLASTX
NCBI GI g166878
BLAST score 154
E value 4.0e-10
Match length 102
% identity 37

NCBI Description (M95796) St12p protein [Arabidopsis thaliana]

Seq. No. 5399

Contig ID 4856_2.R1040

5'-most EST LIB3\overline{1}38-014-Q1-N2-H8

Method BLASTX
NCBI GI g3785976
BLAST score 206
E value 4.0e-16
Match length 128
% identity 37

NCBI Description (AC005560) Sec12p-like protein [Arabidopsis thaliana]

Seq. No. 5400

Contig ID 4857_1.R1040 5'-most EST sat701011352.h1

Method BLASTX
NCBI GI g3548802
BLAST score 416
E value 2.0e-40
Match length 168
% identity 50

NCBI Description (AC005313) axi 1-like protein [Arabidopsis thaliana]

>gi_4335769_gb_AAD17446_ (AC006284) putative axi1 protein

[Nicotiana tabacum] [Arabidopsis thaliana]

Seq. No. 5401

Contig ID 4859_1.R1040

5'-most EST LIB3093-031-Q1-K1-F6

Method BLASTX
NCBI GI g2244924
BLAST score 324

E value 4.0e-30 Match length 81 73

NCBI Description (Z97339) glutaredoxin [Arabidopsis thaliana]

Seq. No. 5402

Contig ID 4859_2.R1040 5'-most EST leu701153189.h1

Method BLASTX
NCBI GI g2244924
BLAST score 226
E value 1.0e-18
Match length 56
% identity 73

NCBI Description (Z97339) glutaredoxin [Arabidopsis thaliana]

Seq. No. 5403

Contig ID 4863_1.R1040 5'-most EST uC-gmropic113h06b1

Method BLASTX
NCBI GI g3776084
BLAST score 280
E value 1.0e-24
Match length 85
% identity 59

NCBI Description (Y18251) NtN2 [Medicago truncatula]

Seq. No. 5404

Contig ID 4863_2.R1040 5'-most EST xzm700763745.h1

Method BLASTX
NCBI GI g3776084
BLAST score 188
E value 1.0e-18
Match length 83
% identity 58

NCBI Description (Y18251) NtN2 [Medicago truncatula]

Seq. No. 5405

Contig ID 4866_1.R1040

5'-most EST LIB3049-015-Q1-E1-B1

Method BLASTX
NCBI GI g2136139
BLAST score 262
E value 2.0e-42
Match length 297
% identity 15

NCBI Description sds22 protein homolog - human >gi_1085028_emb_CAA90626_

(Z50749) yeast sds22 homolog [Homo sapiens]

>gi 1585165_prf__2124310A sds22 gene [Homo sapiens]

>gi_4506013_ref_NP_002703.1_pPPPP1R7_ protein phosphatase 1,

regulatory subunit

Seq. No. 5406

Contig ID 4867_1.R1040

5'-most EST jC-gmst02400009d05d1

Method BLASTX



```
g2673912
NCBI GI
                  510
BLAST score
                   2.0e-51
E value
                  159
Match length
% identity
                   63
NCBI Description (AC002561) unknown protein [Arabidopsis thaliana]
                   5407
Seq. No.
                   4867 2.R1040
Contig ID
                  wvk700685447.h1
5'-most EST
                   BLASTX
Method
                   g2492530
NCBI GI
                   2050
BLAST score
                   0.0e + 00
E value
                   531
Match length
                   74
% identity
                   CHLOROPLAST AMINOPEPTIDASE 1 PRECURSOR (LEUCINE
NCBI Description
                   AMINOPEPTIDASE) (LAP) (LEUCYL AMINOPEPTIDASE) (PROLINE
                   AMINOPEPTIDASE) (PROLYL AMINOPEPTIDASE) >gi_924630 (U20594)
                   leucine aminopeptidase [Solanum lycopersicum]
Seq. No.
                   5408
                   4867 3.R1040
Contig ID
                   uC-gmropic027e08b1
5'-most EST
Method
                   BLASTX
                   g2492530
NCBI GI
                   588
BLAST score
                   1.0e-60
E value
                   196
Match length
                   61
% identity
                   CHLOROPLAST AMINOPEPTIDASE 1 PRECURSOR (LEUCINE
NCBI Description
                   AMINOPEPTIDASE) (LAP) (LEUCYL AMINOPEPTIDASE) (PROLINE
                   AMINOPEPTIDASE) (PROLYL AMINOPEPTIDASE) >gi_924630 (U20594)
                   leucine aminopeptidase [Solanum lycopersicum]
                   5409
Seq. No.
                   4867 4.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy284f02b1
                   5410
Seq. No.
                   4867 5.R1040
Contig ID
                   jC-qmle01810068e01a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1483563
                   485
BLAST score
                   8.0e-49
E value
                   113
Match length
                   77
 % identity
                   (X99825) leucine aminopeptidase [Petroselinum crispum]
NCBI Description
 Seq. No.
                   5411
                   4867_6.R1040
 Contig ID
                   uC-gmropic070c01b1
 5'-most EST
```

BLASTX

421 2.0e-41

g2673912

Method

NCBI GI BLAST score

E value



Match length 106 % identity 75

NCBI Description (AC002561) unknown protein [Arabidopsis thaliana]

Seq. No. 5412

Contig ID 4873_1.R1040

5'-most EST LIB3051-043-Q1-K1-G1

Method BLASTX
NCBI GI g1408460
BLAST score 2168
E value 0.0e+00
Match length 514
% identity 81

NCBI Description (U40161) type 2A protein serine/threonine phosphatase 55

kDa B regulatory subunit [Arabidopsis thaliana]

Seq. No. 5413

Contig ID 4873_2.R1040

5'-most EST jC-gmle01810087b01d1

Seq. No. 5414

Contig ID 4873_3.R1040

5'-most EST jC-gmro02910046a02d1

Method BLASTX
NCBI GI g1408460
BLAST score 341
E value 4.0e-32
Match length 76
% identity 88

NCBI Description (U40161) type 2A protein serine/threonine phosphatase 55

kDa B regulatory subunit [Arabidopsis thaliana]

Seq. No. 5415

Contig ID 4873 4.R1040

5'-most EST LIB3028-009-Q1-B1-F11

Method BLASTX
NCBI GI g3421413
BLAST score 462
E value 2.0e-46
Match length 125
% identity 73

NCBI Description (AF081922) protein phosphatase 2A 55 kDa B regulatory subunit [Oryza sativa] >gi 3421415 (AF081923) protein

phosphatase 2A 55 kDa B regulatory subunit [Oryza sativa]

Seq. No. 5416

Contig ID 4875_1.R1040 5'-most EST vzy700750750.h1

Method BLASTX
NCBI GI g3334323
BLAST score 966
E value 1.0e-105
Match length 193
% identity 93
NCBI Description GTP-BINDI

GTP-BINDING PROTEIN SAR1A >gi_1314860 (U56929) Sar1 homolog

[Arabidopsis thaliana] >gi_2104532_gb_AAC78700.1

(AF001308) SAR1/GTP-binding secretory factor [Arabidopsis



thaliana] >gi 2104550 (AF001535) AGAA.4 [Arabidopsis thaliana]

5417 Seq. No.

4875 2.R1040 Contig ID $asn7\overline{0}1138814.h1$ 5'-most EST

BLASTX Method NCBI GI g3334323 753 BLAST score 8.0e-99 E value 193 Match length 87 % identity

GTP-BINDING PROTEIN SAR1A >gi 1314860 (U56929) Sar1 homolog NCBI Description

[Arabidopsis thaliana] >gi_2104532_gb_AAC78700.1

(AF001308) SAR1/GTP-binding secretory factor [Arabidopsis

thaliana] >gi_2104550 (AF001535) AGAA.4 [Arabidopsis

thaliana]

Seq. No. 5418

4875 3.R1040 Contig ID ncj700982523.hl 5'-most EST

BLASTN Method q2108346 NCBI GI 82 BLAST score 4.0e-38 E value 174 Match length 87 % identity

Brassica campestris small GTP-binding protein Bsar1b NCBI Description

(bsar1b) mRNA, complete cds

5419 Seq. No.

4876 1.R1040 Contig ID ncj700981385.hl 5'-most EST

BLASTX Method q2194137 NCBI GI 321 BLAST score 4.0e-29 E value 99 Match length % identity

(AC002062) ESTs gb R29947,gb_H76702 come from this gene. NCBI Description

[Arabidopsis thaliana]

5420 Seq. No.

Contig ID 4876 2.R1040 ncj700980838.h1 5'-most EST

5421 Seq. No.

4876 3.R1040 Contig ID

LIB3138-025-Q1-N1-B10 5'-most EST

BLASTX Method g2194137 NCBI GI 260 BLAST score E value 4.0e-40Match length 126 75

% identity (AC002062) ESTs gb_R29947,gb_H76702 come from this gene. NCBI Description

[Arabidopsis thaliana]



Seq. No. 5422

Contig ID 4876_4.R1040 5'-most EST epx701106549.h1

Method BLASTX
NCBI GI g2194137
BLAST score 410
E value 5.0e-40
Match length 113
% identity 74

NCBI Description (AC002062) ESTs gb_R29947,gb_H76702 come from this gene.

[Arabidopsis thaliana]

Seq. No. 5423

Contig ID 4878_1.R1040 5'-most EST jsh701066075.h1

Method BLASTX
NCBI GI g4056432
BLAST score 1220
E value 1.0e-134
Match length 448
% identity 60

NCBI Description (AC005990) Similar to gi_2245014 glucosyltransferase

homolog from Arabidopsis thaliana chromosome 4 contig gb Z97341. ESTs gb_T20778 and gb_AA586281 come from this

gene. [Arabidopsis Thaliana]

Seq. No. 5424

Contig ID 4878_2.R1040 5'-most EST seb700652962.h1

Seq. No. 5425

Contig ID 4880_1.R1040

5'-most EST LIB3053-006-Q1-N1-G1

Method BLASTX
NCBI GI g543721
BLAST score 169
E value 1.0e-11
Match length 142
% identity 32

NCBI Description PROTEIN PHOSPHATASE PP2A, 72 KD REGULATORY SUBUNIT (PR72) >gi 539659 pir A47114 phosphoprotein phosphatase (EC

3.1.3.16) 2A regulatory chain PR72 - human >gi_190222 (L12146) protein phosphatase 2A 72 kDa regulatory subunit [Homo sapiens] >gi_4506021_ref_NP_002709.1_pPPP2R3_ protein phosphatase 2 (formerly 2A), regulatory subunit B'' (PR

72), alpha isoform and (PR 130), beta isoform

Seq. No. 5426

Contig ID 4880 2.R1040

5'-most EST uC-gmflminsoy059f03b1

Seq. No. 5427

Contig ID 4886 1.R1040

5'-most EST uC-gmflminsoy058g03b1

Method BLASTX NCBI GI g4126403

1036



```
BLAST score 901
E value 4.0e-97
Match length 331
% identity 52
NCBI Description (AB011796) flavonol synthase [Citrus unshiu]
Seq. No. 5428
Contig ID 4889 1.R1040
```

LIB3087-007-Q1-K1-G9

Seq. No. 5429

5'-most EST

Contig ID 4890 1.R1040 5'-most EST fde700874865.h1

Method BLASTX
NCBI GI g132717
BLAST score 250
E value 3.0e-21
Match length 92
% identity 53

NCBI Description 50S RIBOSOMAL PROTEIN L17 (BL21) >gi_71253_pir__R5BS17 ribosomal protein L17 - Bacillus stearothermophilus

 Seq. No.
 5430

 Contig ID
 4890_2.R1040

 5'-most EST
 bth700845664.h1

 Method
 BLASTX

 NCBI GI
 g3925782

NCBI GI g3925782
BLAST score 138
E value 1.0e-08
Match length 49
% identity 57

NCBI Description (AL034353) putative 60s ribosomal protein

[Schizosaccharomyces pombe]

Seq. No. 5431

Contig ID 4892_1.R1040 5'-most EST dpv701102209.h1

Method BLASTX
NCBI GI g2281102
BLAST score 287
E value 2.0e-25
Match length 139
% identity 52

NCBI Description (AC002333) SF16 isolog [Arabidopsis thaliana]

Seq. No. 5432

Contig ID 4894 1.R1040

5'-most EST LIB3039-005-Q1-E1-G11

Method BLASTN
NCBI GI g562281
BLAST score 123
E value 1.0e-62
Match length 247
% identity 87

NCBI Description B.oleracea mRNA for PSST subunit of NADH:ubiquinone

oxidoreductase



Seq. No. 5433

Contig ID 4894_2.R1040

5'-most EST LIB3170-009-Q2-K2-B8

Method BLASTN
NCBI GI g562281
BLAST score 72
E value 3.0e-32
Match length 168
% identity 86

NCBI Description B.oleracea mRNA for PSST subunit of NADH:ubiquinone

oxidoreductase

Seq. No. 5434

Contig ID 4895_1.R1040

5'-most EST LIB3039-033-Q1-E1-C8

Method BLASTX
NCBI GI g3193303
BLAST score 261
E value 2.0e-22
Match length 113
% identity 46

NCBI Description (AF069298) similar to several proteins containing a tandem

repeat region such as Plasmodium falciparum GGM tandem repeat protein (GB:U27807); partial CDS [Arabidopsis

thaliana]

Seq. No. 5435

Contig ID 4896_1.R1040 5'-most EST zhf700961274.h1

Method BLASTX
NCBI GI 94454042
BLAST score 111
E value 1.0e-08
Match length 155
% identity 34

NCBI Description (AL035394) putative protein [Arabidopsis thaliana]

Seq. No. 5436

Contig ID 4896_2.R1040

5'-most EST LIB3138-040-Q1-N1-D3

Method BLASTX
NCBI GI g2501555
BLAST score 226
E value 2.0e-18
Match length 140
% identity 41

NCBI Description POSSIBLE APOSPORY-ASSOCIATED PROTEIN C >gi_549984 (U13148)

possible apospory-associated protein [Pennisetum ciliare]

Seq. No. 5437

Contig ID 4896_3.R1040

5'-most EST LIB3109-003-Q1-K1-G7

Seq. No. 5438

Contig ID 4899 1.R1040

5'-most EST uC-gmrominsoy247f03b1



```
5439
Seq. No.
                  4899 2.R1040
Contig ID
                  uC-gmronoir009e05b1
5'-most EST
                  BLASTX
Method
                  g130359
NCBI GI
                  201
BLAST score
                  9.0e-16
E value
Match length
                  75
% identity
                   51
                  70 KD PEROXISOMAL MEMBRANE PROTEIN (PMP70)
NCBI Description
                  >gi 111319_pir__A35723 70K peroxisomal membrane protein -
                  rat >gi 220862 dbj_BAA14086_ (D90038) PMP70 [Rattus
                  norvegicus]
                   5440
Seq. No.
                   4899 3.R1040
Contig ID
                   LIB3028-009-Q1-B1-D3
5'-most EST
Seq. No.
                   5441
                   4900 1.R1040
Contig ID
                   LIB3107-010-Q1-K1-B6
5'-most EST
Method
                   BLASTX
                   q4538967
NCBI GI
                   237
BLAST score
E value
                   9.0e-20
                   63
Match length
% identity
                   39
                   (AL049488) major intrinsic protein (MIP)-like [Arabidopsis
NCBI Description
                   thaliana]
                   5442
Seq. No.
                   4902 1.R1040
Contig ID
                   LIB3051-088-Q1-K1-D3
5'-most EST
                   BLASTX
Method
                   q2984196
NCBI GI
                   259
BLAST score
                   4.0e-22
E value
                   140
Match length
                   41
% identity
                   (AE000764) ribosomal protein L11 [Aquifex aeolicus]
NCBI Description
                   5443
Seq. No.
                   4905 1.R1040
Contig ID
5'-most EST
                   zzp700835052.hl
                   BLASTN
Method
                   q3168839
NCBI GI
                   73
BLAST score
                   2.0e-32
E value
Match length
                   145
```

% identity

Arabidopsis thaliana copper homeostasis factor (CCH) mRNA, NCBI Description

complete cds

Seq. No. 5444

4905 2.R1040 Contig ID

LIB3074-012-Q1-E1-H4 5'-most EST

BLASTN Method

NCBI Description



```
g3168839
NCBI GI
                  71
BLAST score
                  1.0e-31
E value
                  187
Match length
                  84
% identity
                  Arabidopsis thaliana copper homeostasis factor (CCH) mRNA,
NCBI Description
                  complete cds
                  5445
Seq. No.
                  4905 3.R1040
Contig ID
                  LIB3167-078-P1-K2-H5
5'-most EST
                  BLASTN
Method
                  g3168839
NCBI GI
                  42
BLAST score
                  5.0e-14
E value
                  62
Match length
                  92
% identity
                  Arabidopsis thaliana copper homeostasis factor (CCH) mRNA,
NCBI Description
                  complete cds
                   5446
Seq. No.
                   4906 1.R1040
Contig ID
                   jex700906490.h1
5'-most EST
                  BLASTX
Method
                   g4469008
NCBI GI
                   177
BLAST score
                   9.0e-13
E value
                   62
Match length
% identity
                  (AL035602) UDP rhamnose--anthocyanidin-3-glucoside
NCBI Description
                   rhamnosyltransferase-like protein [Arabidopsis thaliana]
                   5447
Seq. No.
                   4908 1.R1040
Contig ID
                   uC-qmflminsoy058h04b1
5'-most EST
                   BLASTX
Method
                   g2262116
NCBI GI
BLAST score
                   281
                   6.0e-25
E value
                   135
Match length
% identity
                   (AC002343) cellulose synthase isolog [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5448
                   4910 1.R1040
Contig ID
                   jC-qmf102220057a11d1
5'-most EST
                   5449
Seq. No.
                   4911 1.R1040
Contig ID
                   LIB3093-033-Q1-K1-H9
5'-most EST
Method
                   BLASTX
                   g2285792
NCBI GI
BLAST score
                   559
                   4.0e-57
E value
                   155
Match length
                   70
% identity
```

(AB004568) cyanase [Arabidopsis thaliana]



>gi_3287503_dbj_BAA31224_ (AB015748) cyanase [Arabidopsis thaliana]

Seq. No. 5450

Contig ID 4912_1.R1040

5'-most EST jC-gmst02400018d09a1

Method BLASTX
NCBI GI g3643085
BLAST score 363
E value 3.0e-34
Match length 191
% identity 41

NCBI Description (AF075580) protein phosphatase-2C; PP2C [Mesembryanthemum

crystallinum]

Seq. No. 5451

Contig ID 4913 1.R1040

5'-most EST uC-gmronoir025d01b1

Method BLASTX
NCBI GI g3335372
BLAST score 409
E value 1.0e-39
Match length 112
% identity 65

NCBI Description (AC003028) putative SRG1 protein [Arabidopsis thaliana]

Seq. No. 5452

Contig ID 4915_1.R1040 5'-most EST zpv700757869.h1

Seq. No. 5453

Contig ID 4915_2.R1040

5'-most EST LIB3093-040-Q1-K1-F10

Seq. No. 5454

Contig ID 4918_1.R1040 5'-most EST g4397125

Seq. No. 5455

Contig ID 4921_1.R1040

5'-most EST LIB3107-059-Q1-K1-B11

Method BLASTN
NCBI GI g168498
BLAST score 119
E value 3.0e-60
Match length 315
% identity 84

NCBI Description Corn histone H4 (H4C13) gene, complete cds

Seq. No. 5456

Contig ID 4926_1.R1040

5'-most EST jC-gmro02910066a04a1

Seq. No. 5457

Contig ID 4926 2.R1040

5'-most EST uC-gmflminsoy108e01b1



```
5458
Seq. No.
                  4929 1.R1040
Contig ID
                  LIB3092-048-Q1-K1-H11
5'-most EST
                  BLASTX
Method
                  g4220481
NCBI GI
BLAST score
                  515
                  3.0e-52
E value
                  140
Match length
                  67
% identity
                  (AC006069) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  4929 2.R1040
Contig ID
                  uC-gmropic019c09b1
5'-most EST
                  BLASTX
Method
                  g4220481
NCBI GI
                  376
BLAST score
                   4.0e-36
E value
                   103
Match length
                   70
% identity
NCBI Description (AC006069) unknown protein [Arabidopsis thaliana]
                   5460
Seq. No.
                   4929 3.R1040
Contig ID
                   uC-gmropic106f02b1
5'-most EST
                   BLASTX
Method
                   g4220481
NCBI GI
                   242
BLAST score
                   2.0e-29
E value
                   139
Match length
                   52
% identity
                   (AC006069) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                   4929 4.R1040
                   uC-gmrominsoy105b11b1
5'-most EST
                   5462
Seq. No.
                   4937 1.R1040
Contig ID
                   dpv701099233.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1418990
BLAST score
                   701
E value
                   6.0e-74
                   203
Match length
                   70
% identity
                   (Z75524) unknown [Lycopersicon esculentum]
NCBI Description
Seq. No.
                   5463
                   4937 2.R1040
Contig ID
                   LIB3107-065-Q1-K1-G11
 5'-most EST
                   BLASTX
Method
NCBI GI
                   g1418990
BLAST score
                   487
                   4.0e-49
E value
                   101
Match length
```

88

% identity



(Z75524) unknown [Lycopersicon esculentum] NCBI Description

5464 Seq. No.

Contig ID 4937 3.R1040

LIB3028-008-Q1-B1-F5 5'-most EST

Method BLASTX q1418990 NCBI GI BLAST score 282 3.0e-44E value 103 Match length % identity

NCBI Description (Z75524) unknown [Lycopersicon esculentum]

5465 Seq. No.

4939 1.R1040 Contig ID 5'-most EST pxt700944780.hl

Method BLASTN NCBI GI q434342 144 BLAST score 8.0e-75 E value 464 Match length % identity 83

NCBI Description A.thaliana (C24) mRNA for S18 ribosomal protein

5466 Seq. No.

4939 2.R1040 Contig ID

LIB3040-048-Q1-E1-H12

5'-most EST BLASTX Method

NCBI GI q464707 BLAST score 677 3.0e-71E value Match length 152 % identity 86

40S RIBOSOMAL PROTEIN S18 >gi 480908 pir S37496 ribosomal NCBI Description

protein S18.A - Arabidopsis thaliana

>gi_405613_emb_CAA80684_ (Z23165) ribosomal protein S18A [Arabidopsis thaliana] >gi_434343_emb_CAA82273_ (Z28701)

S18 ribosomal protein [Arabidopsis thaliana]

>gi_434345_emb_CAA82274_ (Z28702) S18 ribosomal protein [Arabidopsis thaliana] >gi_434906_emb_CAA82275_ (Z28962)

S18 ribosomal protein [Arabidopsis thaliana]

>gi_2505871_emb_CAA72909_ (Y12227) ribosomal protein S18A [Arabidopsis thaliana] >gi 3287678 (AC003979) Match to ribosomal S18 gene mRNA gb Z28701, DNA gb Z23165 from A. thaliana. ESTs gb_T21121, gb_Z17755, gb_R64776 and gb_R30430 come from this gene. [Arabidopsis thaliana] >gi_4538910_emb_CAB39647.1_ (AL049482) S18.A ribosomal

protein [Arabidopsis thaliana]

5467 Seq. No.

4939 3.R1040 Contig ID 5'-most EST ncj700983777.h1

BLASTX Method g3746064 NCBI GI BLAST score 146 2.0e-09 E value Match length 80

```
% identity
NCBI Description (AC005311) unknown protein [Arabidopsis thaliana]
                  5468
Seq. No.
                  4939 6.R1040
Contig ID
5'-most EST
                  leu701150769.h1
Method
                  BLASTN
                  q434344
NCBI GI
BLAST score
                  76
                  8.0e-35
E value
                  172
Match length
                  86
% identity
NCBI Description A.thaliana (Columbia) mRNA for S18 ribosomal protein
                  (641bp)
Seq. No.
                  5469
Contig ID
                  4940 1.R1040
                  jex700904203.h1
5'-most EST
                  BLASTX
Method
                  g2961384
NCBI GI
BLAST score
                  612
                  1.0e-63
E value
Match length
                  189
                   59
% identity
NCBI Description (AL022141) aldehyde dehydrogenase like protein [Arabidopsis
                  thaliana]
                   5470
Seq. No.
                   4941 1.R1040
Contig ID
5'-most EST
                  LIB3051-071-Q1-K1-G8
                   5471
Seq. No.
Contig ID
                   4943 1.R1040
5'-most EST
                   LIB3028-008-Q1-B1-E4
                   BLASTX
Method
NCBI GI
                   g1707480
BLAST score
                   195
E value
                   2.0e-29
Match length
                   153
% identity
                   45
                  (Y08614) CRM1 [Homo sapiens]
NCBI Description
Seq. No.
                   5472
                   4944 1.R1040
Contig ID
                   LIB3170-050-Q1-J1-E7
5'-most EST
Method
                   BLASTX
                   g2244906
NCBI GI
BLAST score
                   257
                   4.0e-22
E value
                   122
Match length
```

NCBI Description

(Z97339) indole-3-acetate beta-glucosyltransferase [Arabidopsis thaliana]

Seq. No. Contig ID

% identity

5473

43

4945 1.R1040

5'-most EST

LIB3049-019-Q1-E1-H6



Method BLASTX
NCBI GI g1871184
BLAST score 157
E value 2.0e-10
Match length 66
% identity 59

NCBI Description (U90439) unknown protein [Arabidopsis thaliana]

Seq. No. 5474

Contig ID 4947_1.R1040

5'-most EST jC-gmro02800035g06a1

Method BLASTX
NCBI GI g2213783
BLAST score 258
E value 3.0e-29
Match length 104
% identity 67

NCBI Description (U89256) Pti5 [Lycopersicon esculentum]

Seq. No. 5475

Contig ID 4949 1.R1040

5'-most EST LIB3028-008-Q1-B1-F12

Seq. No. 5476

Contig ID 4951 1.R1040

5'-most EST LIB3094-069-Q1-K1-H3

Method BLASTX
NCBI GI g20020
BLAST score 370
E value 2.0e-35
Match length 83
% identity 92

NCBI Description (X62368) ribosomal protein L12-la [Nicotiana tabacum]

Seq. No. 5477

Contig ID 4951_2.R1040

5'-most EST LIB3106-098-Q1-K1-D4

Method BLASTX
NCBI GI g548655
BLAST score 215
E value 4.0e-17
Match length 137
% identity 39

NCBI Description 50S RIBOSOMAL PROTEIN L12-C, CHLOROPLAST PRECURSOR (CL12-C)

>gi_541897_pir__C53394 ribosomal protein L12.C, chloroplast
- Arabidopsis thaliana >gi 468773_emb_CAA48183_ (X68046)

ribosomal protein L12 [Arabidopsis thaliana]

Seq. No. 5478

Contig ID 4952 1.R1040

5'-most EST uC-gmropic087e09b1

Seq. No. 5479

Contig ID 4957 1.R1040 5'-most EST leu701147618.h1

Method BLASTX NCBI GI g1050849



BLAST score 242 E value 1.0e-19 Match length 210 % identity 32

NCBI Description (X83742) MAP kinase phosphatase [Xenopus laevis]

Seq. No. 5480

Contig ID 4958_1.R1040

5'-most EST LIB3170-069-Q1-J1-D12

Seq. No. 5481

Contig ID 4958_2.R1040

5'-most EST LIB3028-008-Q1-B1-C9

Method BLASTN
NCBI GI g1150683
BLAST score 592
E value 0.0e+00
Match length 720
% identity 97

NCBI Description V.radiata atpB, rbcL and trnK genes

Seq. No. 5482

Contig ID 4959_1.R1040

5'-most EST jC-gmf102220067g01a1

Method BLASTX
NCBI GI g2829863
BLAST score 459
E value 9.0e-56
Match length 211
% identity 56

NCBI Description (AC002396) Putative peroxidase [Arabidopsis thaliana]

Seq. No. 5483

Contig ID 4961_1.R1040 5'-most EST xpa700795669.h1

Method BLASTX
NCBI GI g1946365
BLAST score 401
E value 9.0e-39
Match length 99
% identity 73

NCBI Description (U93215) glutaredoxin isolog [Arabidopsis thaliana]

Seq. No. 5484

Contig ID 4961 2.R1040

5'-most EST LIB3065-009-Q1-N1-A10

Method BLASTX
NCBI GI g1946365
BLAST score 404
E value 3.0e-39
Match length 99
% identity 74

NCBI Description (U93215) glutaredoxin isolog [Arabidopsis thaliana]

Seq. No. 5485

Contig ID 4961_3.R1040 5'-most EST wrg700786730.h2 

Seq. No. 5486

Contig ID 4962 1.R1040

5'-most EST LIB3109-022-Q1-K1-A4

Method BLASTX
NCBI GI g3063706
BLAST score 467
E value 2.0e-46
Match length 131
% identity 65

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 5487

Contig ID 4964 1.R1040

Method BLASTN
NCBI GI g2565418
BLAST score 153
E value 2.0e-80
Match length 385
% identity 85

NCBI Description Onobrychis viciifolia histone H3 mRNA, complete cds

Seq. No. 5488

Contig ID 4964 2.R1040 5'-most EST leu701155595.h1

Method BLASTN
NCBI GI g166383
BLAST score 104
E value 2.0e-51
Match length 240
% identity 86

NCBI Description Alfalfa histone H3 (H3-1.1) gene, complete cds

Seq. No. 5489

Seq. No. 5490

Contig ID 4965_2.R1040 5'-most EST uC-gmropic020h10b1

Seq. No. 5491

Contig ID 4967 1.R1040

5'-most EST LIB3028-008-Q1-B1-A9

Method BLASTN
NCBI GI g11571
BLAST score 380
E value 0.0e+00
Match length 401
% identity 98

NCBI Description Soybean chloroplast rps12 and rps7 genes

Seq. No. 5492

Contig ID 4968 1.R1040

5'-most EST LIB3029-011-Q1-B1-C12

Method BLASTX



```
g3241945
NCBI GI
                  437
BLAST score
                  5.0e-43
E value
                  132
Match length
                   58
% identity
                  (AC004625) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   4968 2.R1040
Contig ID
                   jex700904823.h1
5'-most EST
                   5494
Seq. No.
                   4969 1.R1040
Contig ID
                   LIB3028-008-Q1-B1-B10
5'-most EST
                   5495
Seq. No.
                   4971 1.R1040
Contig ID
                   wvk700682772.h1
5'-most EST
                   BLASTX
Method
                   g4469012
NCBI GI
                   311
BLAST score
                   1.0e-28
E value
                   88
Match length
                   76
% identity
                  (AL035602) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5496
Seq. No.
                   4972 1.R1040
Contig ID
                   jC-gmf102220106a08a1
5'-most EST
                   BLASTX
Method
                   q2804278
NCBI GI
                   1861
BLAST score
                   0.0e + 00
E value
                   414
Match length
% identity
                   (AB003516) squalene epoxidase [Panax ginseng]
NCBI Description
                   5497
Seq. No.
                   4972 2.R1040
Contig ID
                   jC-gmf102220086d06a1
 5'-most EST
                   BLASTX
Method
                   q2804278
NCBI GI
BLAST score
                   698
                   2.0e-75
E value
                   179
Match length
                   79
 % identity
                   (AB003516) squalene epoxidase [Panax ginseng]
 NCBI Description
                    5498
 Seq. No.
                    4973 1.R1040
 Contig ID
                    gsv701050692.hl
 5'-most EST
                    BLASTX
 Method
                    g4262228
 NCBI GI
                    424
 BLAST score
                    7.0e-53
 E value
```

1048

161

67

Match length

% identity



(AC006200) putative receptor protein kinase [Arabidopsis NCBI Description thaliana]

5499 Seq. No.

4974 1.R1040 Contig ID

LIB3040-003-Q1-E1-F10 5'-most EST

5500 Seq. No.

4975 1.R1040 Contig ID 5'-most EST epx701105965.hl

5501 Seq. No.

4975 2.R1040 Contig ID

LIB3092-029-Q1-K1-F1 5'-most EST

5502 Seq. No.

4980 1.R1040 Contig ID kl1701214527.h1 5'-most EST

5503 Seq. No.

4980 2.R1040 Contig ID

LIB3139-047-P1-N1-D1 5'-most EST

BLASTX Method q729486 NCBI GI 180 BLAST score 4.0e-13 E value 102 Match length % identity 41

PRE-RRNA PROCESSING PROTEIN FHL1 >gi_626950_pir__S43738 NCBI Description

transcription activator FHL1 - yeast (Saccharomyces cerevisiae) >gi_454255 emb_CAA82202 (Z28348) Fhl1p [Saccharomyces cerevisiae] >gi_914975 (U32445) Pre-rRNA processing protein Fhl1 (Swiss Prot. accession number

P39521) [Saccharomyces cerevisiae]

5504 Seq. No.

4983 1.R1040 Contig ID

jC-gmf102220079h05a1 5'-most EST

BLASTX Method g577301 NCBI GI BLAST score 188 E value 6.0e-14Match length 108 % identity 34

(D42044) The ha3523 gene product is related to S.cerevisiae NCBI Description

gene product located in chromosome III. [Homo sapiens]

5505 Seq. No.

4985 1.R1040 Contig ID zpv700758101.h1 5'-most EST

BLASTX Method q4510425 NCBI GI BLAST score 620 2.0e-64 E value 140 Match length 83 % identity

NCBI Description (AC006929) unknown protein [Arabidopsis thaliana]



```
5506
Seq. No.
                  4986 1.R1040
Contig ID
                  jex700907662.h1
5'-most EST
                  BLASTX
Method
                  g1458245
NCBI GI
                  397
BLAST score
                  4.0e-38
E value
                  254
Match length
                  34
% identity
                  (U64834) coded for by C. elegans cDNA cml7al; coded for by
NCBI Description
                  C. elegans cDNA cm7g1; coded for by C. elegans cDNA
                  CEMSE26F; similar to methyltransferases [Caenorhabditis
                  elegans]
                   5507
Seq. No.
                   4987 1.R1040
Contig ID
                   awf700838030.h1
5'-most EST
                  BLASTX
Method
                   g2623299
NCBI GI
                   306
BLAST score
                   5.0e-28
E value
                   127
Match length
                   50
% identity
NCBI Description (AC002409) hypothetical protein [Arabidopsis thaliana]
                   5508
Seq. No.
                   4989 1.R1040
Contig ID
                   LIB3139-043-P1-N1-D2
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2191129
BLAST score
                   489
E value
                   6.0e-49
                   129
Match length
% identity
                   (AF007269) similar to SPF1 DNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
                   5509
Seq. No.
                   4997 1.R1040
Contig ID
                   leu701145678.hl
5'-most EST
Seq. No.
                   5510
                   4997 2.R1040
Contig ID
                   asn701136067.h1
5'-most EST
                   5511
Seq. No.
Contig ID
                   4999 1.R1040
                   LIB3028-007-Q1-B1-F6
 5'-most EST
Method
                   BLASTX
                   g544110
NCBI GI
BLAST score
                   212
E value
                   5.0e-17
Match length
                   59
```

1050

COLD SHOCK-LIKE PROTEIN CSPE >gi 1363394_pir _S49050 CspE

protein - Escherichia coli >gi_471099_dbj_BAA05856_

68

37774×

% identity

NCBI Description



(D28497) CspE (MsmC) [Escherichia coli] >gi_833769 (L29054) gicA gene product [Escherichia coli] >gi_1651256_dbj_BAA35266_ (D90703) CspE protein [Escherichia coli] >gi_1786841 (AE000167) cold shock protein [Escherichia coli] >gi_3851642 (AF098260) unknown [Vibrio cholerae]

Seq. No. 5512

Contig ID 5000_1.R1040

5'-most EST LIB3028-007-Q1-B1-F8

Seq. No. 5513

Contig ID 5002_1.R1040

5'-most EST LIB3028-007-Q1-B1-G10

Method BLASTX
NCBI GI g2347195
BLAST score 198
E value 3.0e-15
Match length 67
% identity 54

NCBI Description (AC002338) hypothetical protein [Arabidopsis thaliana]

Seq. No. 5514

Contig ID 5003_1.R1040 5'-most EST bth700846021.h1

Method BLASTX
NCBI GI g3790581
BLAST score 316
E value 1.0e-28
Match length 175
% identity 39

NCBI Description (AF079179) RING-H2 finger protein RHB1a [Arabidopsis

thaliana]

Seq. No. 5515

Contig ID 5003_2.R1040

5'-most EST LIB3028-007-Q1-B1-G2

Seq. No. 5516

Contig ID 5006 1.R1040

5'-most EST LIB3028-007-Q1-B1-G7

Seq. No. 5517

Contig ID 5007 1.R1040

5'-most EST LIB3028-007-Q1-B1-G9

Method BLASTX
NCBI GI g1888357
BLAST score 497
E value 4.0e-50
Match length 189
% identity 54

NCBI Description (X98130) alpha-mannosidase [Arabidopsis thaliana]

>qi 1890154 emb CAA72432 (Y11767) alpha-mannosidase

precursor [Arabidopsis thaliana]

Seq. No. 5518

Contig ID 5008_1.R1040

% identity

NCBI Description

36



```
5'-most EST
                   fua701039005.hl
Method
                   BLASTX
NCBI GI
                   q3560264
BLAST score
                   367
E value
                   1.0e-34
                   219
Match length
% identity
                   42
NCBI Description
                   (AL031535) soll family protein [Schizosaccharomyces pombe]
Seq. No.
                   5519
Contig ID
                   5008 2.R1040
5'-most EST
                   hyd700729250.h1
Method
                   BLASTX
NCBI GI
                   q731737
BLAST score
                   155
E value
                   5.0e-10
                   129
Match length
                   33
% identity
                   SOL3 PROTEIN >gi_626676_pir__S48903 SOL3 protein - yeast
NCBI Description
                   (Saccharomyces cerevisiae) >gi_458904 (U00027) Sol3p [Saccharomyces cerevisiae] >gi_1184943 (U46560) Sol3p
                   [Saccharomyces cerevisiae]
Seq. No.
                   5520
Contig ID
                   5009 1.R1040
5'-most EST
                   jC-gmle01810043g06d1
Method
                   BLASTX
                   q1653626
NCBI GI
BLAST score
                   182
E value
                   3.0e-13
Match length
                   84
% identity
                   44
                   (D90915) hypothetical protein [Synechocystis sp.]
NCBI Description
                   5521
Seq. No.
                   5010 1.R1040
Contig ID
5'-most EST
                   epx701106413.h1
Method
                   BLASTX
                   q4455787
NCBI GI
BLAST score
                   332
                   8.0e-31
E value
Match length
                   147
                   43
% identity
NCBI Description
                   (AL035536) putative DNA polymerase accessory protein
                    [Schizosaccharomyces pombe]
Seq. No.
                   5522
                   5011 1.R1040
Contig ID
5'-most EST
                   LIB3028-007-Q1-B1-D8
Method
                   BLASTX
NCBI GI
                   g4263710
BLAST score
                   455
E value
                   3.0e-45
Match length
                   173
```

protein [Arabidopsis thaliana]

(AC006223) putative pur-alpha transcriptional activator



Seq. No. 5523

Contig ID 5012_1.R1040 5'-most EST taw700660051.h1

Method BLASTX
NCBI GI g400198
BLAST score 199
E value 3.0e-15
Match length 101
% identity 50

NCBI Description PHOTOSYSTEM II 5 KD PROTEIN PRECURSOR (PSII-T)

(LIGHT-REGULATED UNKNOWN 11 KD PROTEIN)

>gi_99604_pir__S21023 hypothetical protein - upland cotton
>gi_1361978_pir__A57500 photosystem II protein psbT upland cotton >gi_18512_emb_CAA38027_ (X54092) unknown

reading frame [Gossypium hirsutum]

Seq. No. 5524

Contig ID 5012_2.R1040

5'-most EST jC-gmst02400063f12a1

Method BLASTX
NCBI GI g400198
BLAST score 190
E value 5.0e-14
Match length 104
% identity 43

NCBI Description PHOTOSYSTEM II 5 KD PROTEIN PRECURSOR (PSII-T)

(LIGHT-REGULATED UNKNOWN 11 KD PROTEIN)

>gi_99604_pir__S21023 hypothetical protein - upland cotton
>gi_1361978_pir__A57500 photosystem II protein psbT upland cotton >gi_18512_emb_CAA38027_ (X54092) unknown

reading frame [Gossypium hirsutum]

Seq. No. 5525

Contig ID 5012_3.R1040

5'-most EST LIB3074-004-Q1-K1-E8

Method BLASTX
NCBI GI g400198
BLAST score 192
E value 2.0e-14
Match length 103
% identity 50

NCBI Description PHOTOSYSTEM II 5 KD PROTEIN PRECURSOR (PSII-T)

(LIGHT-REGULATED UNKNOWN 11 KD PROTEIN)

>gi_99604_pir__S21023 hypothetical protein - upland cotton
>gi_1361978_pir__A57500 photosystem II protein psbT upland cotton >gi_18512_emb_CAA38027_ (X54092) unknown

reading frame [Gossypium hirsutum]

Seq. No. 5526

Contig ID 5013_1.R1040
5'-most EST g4313283
Method BLASTX
NCBI GI g4262250
BLAST score 1690
E value 0.0e+00
Match length 382



% identity 86

NCBI Description (AC006200) putative aldolase [Arabidopsis thaliana]

Seq. No. 5527

Contig ID 5013_2.R1040 5'-most EST uC-gmropic010g08b1

Method BLASTX
NCBI GI 94262250
BLAST score 502
E value 9.0e-51
Match length 145
% identity 74

NCBI Description (AC006200) putative aldolase [Arabidopsis thaliana]

Seq. No. 5528

Contig ID 5013_3.R1040 5'-most EST seb700653622.h1

Seq. No. 5529

Contig ID 5018_1.R1040 5'-most EST uaw700666311.h1

Seq. No. 5530

Contig ID 5021_1.R1040

5'-most EST jC-gmro02910014e08a1

Method BLASTX
NCBI GI g1769887
BLAST score 2015
E value 0.0e+00
Match length 471
% identity 78

NCBI Description (X95736) amino acid permease 6 [Arabidopsis thaliana]

Seq. No. 5531

Contig ID 5021_3.R1040 5'-most EST rca700996181.h1

Method BLASTX
NCBI GI g1769887
BLAST score 292
E value 2.0e-26
Match length 72
% identity 85

NCBI Description (X95736) amino acid permease 6 [Arabidopsis thaliana]

Seq. No. 5532

Contig ID 5024 1.R1040 5'-most EST seb700651789.h1

Method BLASTX
NCBI GI g3128168
BLAST score 1075
E value 1.0e-117
Match length 253
% identity 72

NCBI Description (AC004521) putative carboxyl-terminal peptidase

[Arabidopsis thaliana]

Seq. No. 5533

1054



Contig ID 5025 1.R1040

5'-most EST uC-gmflminsoy118b03b1

Method BLASTX
NCBI GI 94567235
BLAST score 150
E value 9.0e-10
Match length 81
% identity 47

NCBI Description (AC007119) putative phosphatidylinositol/phophatidylcholine

transfer protein [Arabidopsis thaliana]

Seq. No. 5534

Contig ID 5026 1.R1040

5'-most EST LIB3093-014-Q1-K1-G6

Seq. No. 5535

Contig ID 5026 2.R1040

5'-most EST jC-gmle01810089h07a1

Seq. No. 5536

Contig ID 5028_1.R1040 5'-most EST gsv701046547.h1

Method BLASTX
NCBI GI g2398829
BLAST score 302
E value 2.0e-27
Match length 72
% identity 39

NCBI Description (Y11220) mitochondrial uncoupling protein [Solanum

tuberosum]

Seq. No. 5537

Contig ID 5029_1.R1040 5'-most EST vzy700753080.h1

Method BLASTX
NCBI GI g1946329
BLAST score 1014
E value 1.0e-110
Match length 278
% identity 73

NCBI Description (U69154) prohibitin [Nicotiana tabacum]

Seq. No. 5538

Contig ID 5029_2.R1040

5'-most EST LIB3167-027-P1-K1-E3

Method BLASTX
NCBI GI g1946329
BLAST score 420
E value 3.0e-41
Match length 111
% identity 72

NCBI Description (U69154) prohibitin [Nicotiana tabacum]

Seq. No. 5539

Contig ID 5033 1.R1040

5'-most EST LIB3138-031-Q1-N1-D11

Method BLASTN



```
g1370147
NCBI GI
                  317
BLAST score
                  1.0e-178
E value
                  529
Match length
                  91
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB11D
Seq. No.
                   5033 2.R1040
Contig ID
                  pmv700891967.hl
5'-most EST
                  BLASTN
Method
                  g1370149
NCBI GI
                   127
BLAST score
                   7.0e-65
E value
                   203
Match length
                   91
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB11E
                   5541
Seq. No.
                   5035 1.R1040
Contig ID
                   ssr700553580.hl
5'-most EST
                   BLASTN
Method
                   g343022
NCBI GI
                   191
BLAST score
                   1.0e-103
E value
                   500
Match length
                   91
% identity
                   Pea chloroplast photosystem II gene encoding the D2 and
NCBI Description
                   44kd chlorophyll a-binding proteins, complete cds
                   5542
Seq. No.
                   5036 1.R1040
Contig ID
5'-most EST
                   LIB3030-007-Q1-B1-D2
                   5543
Seq. No.
                   5041 1.R1040
Contig ID
                   LIB3028-007-Q1-B1-A5
 5'-most EST
                   BLASTX
Method
                   g1143511
NCBI GI
                   1630
BLAST score
                   0.0e + 00
E value
                   303
Match length
 % identity
                    (Z47076) Ser/Thr protein phosphatase homologous to PPX
 NCBI Description
                    [Malus domestica] >gi_1586034_prf__2202340A Ser/Thr protein
                   phosphatase [Malus domestica]
 Seq. No.
                    5041 3.R1040
 Contig ID
                    LIB3170-019-Q1-K1-H10
 5'-most EST
                    BLASTN
 Method
                    g1143510
 NCBI GI
                    154
 BLAST score
 E value
                    4.0e-81
 Match length
                    350
 % identity
                    86
 NCBI Description M.domestica Borkh mRNA for serine/threonine protein
```



phosphatase (PPX)

Seq. No. 5545
Contig ID 5041 5.R1040
5'-most EST LIB3170-059-Q1-K1-A7
Method BLASTN

NCBI GI g1143510
BLAST score 139
E value 3.0e-72
Match length 291
% identity 87

NCBI Description M.domestica Borkh mRNA for serine/threonine protein

phosphatase (PPX)

Seq. No. 5546

Contig ID 5041_6.R1040 5'-most EST pcp700989770.h1

Method BLASTN
NCBI GI g1143510
BLAST score 67
E value 2.0e-29
Match length 131
% identity 88

NCBI Description M.domestica Borkh mRNA for serine/threonine protein

phosphatase (PPX)

Seq. No. 5547

Contig ID 5042_1.R1040

5'-most EST uC-gmropic090g05b1

Method BLASTX
NCBI GI g1881585
BLAST score 361
E value 6.0e-34
Match length 124

% identity 58

NCBI Description (U72489) remorin [Solanum tuberosum]

Seq. No. 5548

Contig ID 5044_1.R1040 5'-most EST ncj700983465.h1

Method BLASTX
NCBI GI g1402878
BLAST score 1075
E value 1.0e-117
Match length 444
% identity 45

NCBI Description (X98130) unknown [Arabidopsis thaliana]

Seq. No. 5549

Contig ID 5044 2.R1040

5'-most EST jC-gmf102220132aa06d1

Method BLASTX
NCBI GI g3367522
BLAST score 272
E value 9.0e-24
Match length 83
% identity 57



NCBI Description (AC004392) EST gb_T04691 comes from this gene. [Arabidopsis thaliana]

Seq. No. 5550

Contig ID 5044 3.R1040

5'-most EST jC-gmle01810042e08d1

Method BLASTX
NCBI GI g3367522
BLAST score 492
E value 2.0e-49
Match length 196
% identity 47

NCBI Description (AC004392) EST gb_T04691 comes from this gene. [Arabidopsis

thaliana]

Seq. No. 5551

Contig ID 5044 4.R1040

5'-most EST LIB3170-066-Q1-K1-H5

Method BLASTX
NCBI GI g1402878
BLAST score 253
E value 1.0e-21
Match length 94
% identity 50

NCBI Description (X98130) unknown [Arabidopsis thaliana]

Seq. No. 5552

Contig ID 5044_5.R1040 5'-most EST zzp700833553.h1

Method BLASTX
NCBI GI g3367522
BLAST score 311
E value 4.0e-28
Match length 186
% identity 45

NCBI Description (AC004392) EST gb_T04691 comes from this gene. [Arabidopsis

thaliana]

Seq. No. 5553

Contig ID 5044_9.R1040 5'-most EST leu701157040.h1

Seq. No. 5554

Contig ID 5045 1.R1040

5'-most EST LIB3138-042-Q1-N1-E1

Seq. No. 5555

Contig ID 5046 2.R1040

5'-most EST LIB3139-064-P1-N1-H8

Seq. No. 5556

Contig ID 5048_1.R1040 5'-most EST asj700967448.h1

Seq. No. 5557

Contig ID 5049_1.R1040

5'-most EST LIB3106-019-Q1-K1-F2



Method BLASTX
NCBI GI g1655653
BLAST score 223
E value 9.0e-18
Match length 149
% identity 36

NCBI Description (Z81368) hypothetical protein Rv2406c [Mycobacterium

tuberculosis]

Seq. No. 5558

Contig ID 5049 2.R1040 5'-most EST g4313995

Seq. No. 5559

Contig ID 5049_4.R1040 5'-most EST awf700843120.h1

Seq. No. 5560

Contig ID 5051 1.R1040

5'-most EST LIB3\overline{10}-\text{Q1-K1-C3}

Method BLASTX
NCBI GI g2828280
BLAST score 1020
E value 1.0e-111
Match length 206
% identity 93

NCBI Description (AL021687) putative protein [Arabidopsis thaliana]

>gi 2832633 emb_CAA16762 (AL021711) putative protein

[Arabidopsis thaliana]

Seq. No. 5561

Contig ID 5051_2.R1040 5'-most EST zzp700835746.h1

Method BLASTX
NCBI GI g2828280
BLAST score 156
E value 2.0e-10
Match length 33
% identity 88

NCBI Description (AL021687) putative protein [Arabidopsis thaliana]

>gi 2832633 emb CAA16762 (AL021711) putative protein

[Arabidopsis thaliana]

Seq. No. 5562

Contig ID 5055_1.R1040

5'-most EST uC-gmflminsoy045h04b1

Method BLASTX
NCBI GI g3860250
BLAST score 1604
E value 1.0e-179
Match length 398
% identity 80

NCBI Description (AC005824) putative chloroplast prephenate dehydratase

[Arabidopsis thaliana]

Seq. No. 5563

Contig ID 5055_2.R1040



```
5'-most EST
                   uC-qmropic111d07b1
Method
                   BLASTN
NCBI GI
                   q2351064
BLAST score
                   35
E value
                   4.0e-10
Match length
                   47
% identity
                   94
NCBI Description
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
                  MDJ22, complete sequence [Arabidopsis thaliana]
                   5564
Seq. No.
                   5057 1.R1040
Contig ID
5'-most EST
                   LIB3028-006-Q1-B1-H10
Method
                  BLASTX
NCBI GI
                   g4559398
BLAST score
                   <del>7</del>96
E value
                   6.0e-85
                   229
Match length
% identity
                   66
                   (AC006526) putative mitochondrial carrier protein
NCBI Description
                   [Arabidopsis thaliana]
                   5565
Seq. No.
Contig ID
                   5060 1.R1040
5'-most EST
                   LIB3139-101-P1-N1-B4
Method
                   BLASTX
NCBI GI
                   q3451068
BLAST score
                   471
E value
                   4.0e-47
Match length
                   159
% identity
                   64
                  (AL031326) putative protein [Arabidopsis thaliana]
NCBI Description
                   5566
Seq. No.
                   5060 2.R1040
Contig ID
5'-most EST
                   LIB3028-008-Q1-B1-F3
Method
                   BLASTX
NCBI GI
                   q3451067
BLAST score
                   306
E value
                   6.0e-28
Match length
                   99
% identity
                   64
                  (AL031326) putative protein [Arabidopsis thaliana]
NCBI Description
                   5567
Seq. No.
                   5060 3.R1040
Contig ID
5'-most EST
                   LIB3106-087-Q1-K1-A3
Method
                   BLASTX
NCBI GI
                   q3451067
```

BLAST score 137 E value 5.0e-15 Match length 82 % identity

NCBI Description (AL031326) putative protein [Arabidopsis thaliana]

Seq. No. 5568

Contig ID 5060 4.R1040



```
5'-most EST
                  LIB3106-076-Q1-K1-G6
                  BLASTX
Method
                  g3451067
NCBI GI
BLAST score
                  373
                  1.0e-35
E value
                  99
Match length
                  75
% identity
NCBI Description
                  (AL031326) putative protein [Arabidopsis thaliana]
Seq. No.
                  5060 5.R1040
Contig ID
                  leu701155251.h1
5'-most EST
                  5570
Seq. No.
                  5062 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy067f01b1
                  BLASTX
Method
NCBI GI
                  g4204267
BLAST score
                  639
                  1.0e-66
E value
                  201
Match length
                  67
% identity
NCBI Description (AC005223) 55585 [Arabidopsis thaliana]
Seq. No.
                  5571
Contig ID
                  5063 1.R1040
                  uC-gmrominsoy045f06b1
5'-most EST
Method
                  BLASTX
                  g1351856
NCBI GI
BLAST score
                  1789
                  0.0e+00
E value
Match length
                  375
                  90
% identity
NCBI Description
                  ACONITATE HYDRATASE, CYTOPLASMIC (CITRATE HYDRO-LYASE)
                   (ACONITASE) >gi_868003_dbj_BAA06108_ (D29629) aconitase
                   [Cucurbita sp.]
                   5572
Seq. No.
                  5063 2.R1040
Contig ID
                  LIB3028-001-Q1-B1-D1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g1351856
BLAST score
                   432
E value
                  2.0e-42
Match length
                  106
                  79
% identity
                  ACONITATE HYDRATASE, CYTOPLASMIC (CITRATE HYDRO-LYASE)
NCBI Description
                   (ACONITASE) >gi 868003 dbj BAA06108 (D29629) aconitase
                   [Cucurbita sp.]
                   5573
Seq. No.
Contig ID
                  5065 1.R1040
```

5'-most EST leu701149686.hl

BLASTN Method NCBI GI q710349 445 BLAST score E value 0.0e + 00



Match length 757 % identity 90

NCBI Description Lupinus albus farnesyl pyrophosphate synthase (fps1) mRNA,

complete cds

Seq. No. 5574

Contig ID 5065 2.R1040 5'-most EST seb700648557.h1

Method BLASTN
NCBI GI g710349
BLAST score 302
E value 1.0e-169
Match length 546
% identity 89

NCBI Description Lupinus albus farnesyl pyrophosphate synthase (fps1) mRNA,

complete cds

Seq. No. 5575

Contig ID 5065 4.R1040

5'-most EST LIB3139-117-P1-N1-H12

Method BLASTN
NCBI GI g558922
BLAST score 216
E value 1.0e-118
Match length 352
% identity 90

NCBI Description Lupinus albus farnesyl pyrophosphate synthase (fps1) mRNA,

complete cds

Seq. No. 5576

Contig ID 5069_1.R1040 5'-most EST zhf700955726.h1

Seq. No. 5577

Contig ID 5073_1.R1040

5'-most EST LIB3039-018-Q1-E1-B5

Method BLASTX
NCBI GI g3860308
BLAST score 357
E value 2.0e-33
Match length 132
% identity 63

NCBI Description (AJ012681) hypothetical protein [Cicer arietinum]

Seq. No. 5578

Contig ID 5073 2.R1040

5'-most EST LIB3039-010-Q1-E1-D4

Method BLASTX
NCBI GI g3860308
BLAST score 375
E value 1.0e-35
Match length 122
% identity 69

NCBI Description (AJ012681) hypothetical protein [Cicer arietinum]

Seq. No. 5579

Contig ID 5077 1.R1040

Match length

% identity

80

47



```
5'-most EST
                  LIB3030-001-Q1-B1-D10
Method
                  BLASTX
NCBI GI
                  g1730328
                  825
BLAST score
                  3.0e-88
E value
Match length
                  210
                  73
% identity
                  (U78100) unknown protein 038 [Phalaenopsis sp. 'hybrid
NCBI Description
                  SM9108']
Seq. No.
                  5580
Contig ID
                  5079 1.R1040
                  LIB3073-020-Q1-K1-D10
5'-most EST
Method
                  BLASTX
NCBI GI
                  q1255954
BLAST score
                  451
E value
                  1.0e-44
Match length
                  106
% identity
                  78
NCBI Description
                  (Z70677) thioredoxin [Ricinus communis]
Seq. No.
                  5581
                  5079 2.R1040
Contig ID
5'-most EST
                  LIB3072-054-Q1-E1-F1
Method
                  BLASTX
NCBI GI
                  g267124
BLAST score
                  424
E value
                  1.0e-41
Match length
                  105
% identity
                  75
NCBI Description
                  THIOREDOXIN H-TYPE 1 (TRX-H1) >qi 100387 pir S16590
                  thioredoxin h1 - common tobacco >qi 20047 emb CAA41415
                   (X58527) thioredoxin [Nicotiana tabacum]
                  5582
Seq. No.
Contig ID
                  5080 1.R1040
5'-most EST
                  trc700567583.h1
Method
                  BLASTX
                  q3123295
NCBI GI
BLAST score
                  408
E value
                  1.0e-39
Match length
                  136
% identity
                  38
                  CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi 2583169
NCBI Description
                   (AF026473) calmodulin-related protein [Arabidopsis
                  thaliana]
Seq. No.
                  5583
                  5083 1.R1040
Contig ID
5'-most EST
                  LIB3138-128-Q1-N1-B11
Method
                  BLASTX
NCBI GI
                  q2494140
BLAST score
                  192
E value
                  2.0e-14
```

NCBI Description (AC003002) R29515 1 [Homo sapiens] >qi 3068727 (AF058918)



5584

unknown [Homo sapiens]

5083 2.R1040 Contig ID jsh701064323.h1 5'-most EST BLASTX Method g2494140 NCBI GI 350 BLAST score E value 1.0e-32 Match length 135 % identity 52

(AC003002) R29515 1 [Homo sapiens] >gi_3068727 (AF058918) NCBI Description

unknown [Homo sapiens]

Seq. No. 5585

Seq. No.

Contig ID 5088 1.R1040

5'-most EST jC-gmst02400076e12a1

Method BLASTX NCBI GI g1362490 BLAST score 284 5.0e-25 E value Match length 159 41 % identity

NCBI Description hypothetical protein YFR024c-a - yeast (Saccharomyces

cerevisiae)

Seq. No.

5089 1.R1040 Contig ID

LIB3139-005-P1-N1-H5 5'-most EST

5586

Method BLASTN g2736287 NCBI GI BLAST score 191 E value 1.0e-103 Match length 427

% identity 86

Camptotheca acuminata isopentenyl diphosphate isomerase II NCBI Description

(IPI2) mRNA, complete cds

5587 Seq. No.

5089 2.R1040 Contig ID

LIB3\overline{1}38-121-Q1-N1-E9 5'-most EST

Method BLASTN q1695182 NCBI GI BLAST score 118 E value 1.0e-59 Match length 210 % identity

NCBI Description N.tabacum mRNA for isopentenyl pyrophosphate isomerase

5588 Seq. No.

5090 1.R1040 Contig ID jex700904292.hl 5'-most EST

Method BLASTX NCBI GI g4522008 BLAST score 307 E value 4.0e-28 Match length 95



% identity 55

NCBI Description (AC007069) hypothetical protein [Arabidopsis thaliana]

Seq. No.

5589

Contig ID

5091 1.R1040

5'-most EST

uC-gmrominsoy178h09b1

Seq. No.

5590

Contig ID

5100_1.R1040

5'-most EST

uC-gmrominsoy099c12b1

Method NCBI GI BLAST score BLASTX q2224901 162

E value Match length

5.0e-11 41 73

NCBI Description (U67134) PcMYB1 protein [Petroselinum crispum]

Seq. No.

% identity

5591

BLASTX

852

q3236237

1.0e-91

Contig ID

5103 1.R1040

5'-most EST

LIB3039-010-Q1-E1-A5

Method NCBI GI BLAST score E value

270 Match length % identity

NCBI Description

(AC004684) putative ribotol dehydrogenase [Arabidopsis

thaliana]

Seq. No.

5592

Contig ID

5104 1.R1040

5'-most EST

LIB3093-016-Q1-K2-E5

Method BLASTX NCBI GI g2281627 BLAST score 370 E value 4.0e-35 Match length 127 % identity 61

NCBI Description

(AF003094) AP2 domain containing protein RAP2.1

[Arabidopsis thaliana]

Seq. No.

5593

Contig ID

5104 2.R1040

5'-most EST

LIB3093-016-Q1-K2-F6

Method BLASTX NCBI GI g2464890 BLAST score 283 E value 5.0e-25 60 Match length 85

% identity NCBI Description

(Z99707) TINY-like protein [Arabidopsis thaliana]

Seq. No.

5594

Contig ID

5113 1.R1040

5'-most EST

LIB3170-083-Q1-K1-H1

Method

BLASTX



```
NCBI GI g2827656
BLAST score 282
E value 6.0e-25
Match length 73
% identity 71
```

NCBI Description (AL021637) DAG-like protein [Arabidopsis thaliana]

Seq. No. 5595

Contig ID 5113_2.R1040

5'-most EST LIB3109-034-Q1-K1-G5

Seq. No. 5596

Contig ID 5113_3.R1040 5'-most EST bth700846831.h1

Seq. No. 5597

Contig ID 5115 1.R1040

5'-most EST LIB3049-044-Q1-E1-H9

Method BLASTX
NCBI GI g130271
BLAST score 543
E value 3.0e-55
Match length 165
% identity 66

NCBI Description PLASTOCYANIN PRECURSOR >gi 100238 pir __S05303 plastocyanin

precursor - tomato >qi 19300 emb CAA32121 (X13934)

pre-plastocyanin (AA -64 to 106) [Lycopersicon esculentum]

Seq. No. 5598

Contig ID 5115 2.R1040 5'-most EST g5677250 Method BLASTX NCBI GI g130265 BLAST score 467 E value 1.0e-46 Match length 99 % identity 88

NCBI Description PLASTOCYANIN >gi 65834 pir CUVM plastocyanin - field

pumpkin

Seq. No.

5599 5115

Contig ID 5115_3.R1040

5'-most EST LIB3040-007-Q1-E1-B6

Method BLASTX
NCBI GI g130271
BLAST score 518
E value 2.0e-52
Match length 167
% identity 64

NCBI Description PLASTOCYANIN PRECURSOR >gi 100238 pir S05303 plastocyanin

precursor - tomato >gi_19300 emb_CAA32121_ (X13934)

pre-plastocyanin (AA -64 to 106) [Lycopersicon esculentum]

Seq. No. 5600

Contig ID 5115 4.R1040

5'-most EST LIB3073-006-Q1-K1-F9

Method BLASTN



```
NCBI GI g20845
BLAST score 34
E value 1.0e-09
Match length 74
% identity 86
NCBI Description Pea gene for plastocyanin
```

Seq. No. 5601

Contig ID 5119_1.R1040 5'-most EST fde700873747.h1

Method BLASTX
NCBI GI g4539423
BLAST score 704
E value 2.0e-74
Match length 164
% identity 82

NCBI Description (AL049171) pyrophosphate-dependent phosphofructo-1-kinase

[Arabidopsis thaliana]

Seq. No. 5602

Contig ID 5120 1.R1040

5'-most EST LIB3053-001-Q1-B1-C5

Method BLASTX
NCBI GI g1352245
BLAST score 512
E value 8.0e-52
Match length 226
% identity 46

NCBI Description FATTY ALDEHYDE DEHYDROGENASE (ALDEHYDE DEHYDROGENASE,

MICROSOMAL) (CLASS 3) >gi 687254 (U14390) aldehyde

deydrogenase [Mus musculus]

Seq. No. 5603

Contig ID 5121_1.R1040 5'-most EST zsg701129386.h1

Method BLASTX
NCBI GI g4325282
BLAST score 233
E value 2.0e-19
Match length 57
% identity 67

NCBI Description (AF123310) NAC domain protein NAM [Arabidopsis thaliana]

>gi_4325286_gb_AAD17314_ (AF123311) NAC domain protein NAM

[Arabidopsis thaliana]

Seq. No. 5604

Contig ID 5121 2.R1040

5'-most EST uC-gmflminsoy029b08b1

Method BLASTX
NCBI GI 94325282
BLAST score 578
E value 1.0e-59
Match length 148
% identity 72

NCBI Description (AF123310) NAC domain protein NAM [Arabidopsis thaliana]

>gi 4325286 gb AAD17314 (AF123311) NAC domain protein NAM

[Arabidopsis thaliana]



Seq. No. 5605

Contig ID 5122_1.R1040

5'-most EST LIB3106-068-P1-K1-E6

Method BLASTN
NCBI GI g3860330
BLAST score 155
E value 2.0e-81
Match length 387
% identity 85

NCBI Description Cicer arietinum mRNA for hypothetical protein, clone Can106

Seq. No. 5606

Contig ID 5122_2.R1040 5'-most EST kl1701207777.h1

Method BLASTN
NCBI GI g3860330
BLAST score 112
E value 5.0e-56
Match length 236
% identity 87

NCBI Description Cicer arietinum mRNA for hypothetical protein, clone Can106

Seq. No. 5607

Contig ID 5132_1.R1040 5'-most EST fua701038545.h1

Seq. No. 5608

Contig ID 5133 1.R1040

5'-most EST LIB3138-057-Q1-N1-D4

Method BLASTX
NCBI GI g2129517
BLAST score 1719
E value 0.0e+00
Match length 359
% identity 86

NCBI Description 24-sterol C-methyltransferase (EC 2.1.1.41) - Arabidopsis

thaliana

Seq. No. 5609

Contig ID 5133_2.R1040

5'-most EST uC-gmropic067d03b1

Method BLASTX
NCBI GI g2129517
BLAST score 557
E value 3.0e-57
Match length 124
% identity 80

NCBI Description 24-sterol C-methyltransferase (EC 2.1.1.41) - Arabidopsis

thaliana

Seq. No. 5610

Contig ID 5133_4.R1040

5'-most EST LIB3092-030-Q1-K1-H5

Method BLASTX
NCBI GI g2129517
BLAST score 333

1068



E value 4.0e-31 Match length 153 76 % identity

NCBI Description 24-sterol C-methyltransferase (EC 2.1.1.41) - Arabidopsis

thaliana

5611 Seq. No.

5134 1.R1040 Contig ID 5'-most EST g5175390 Method BLASTX NCBI GI q464981 BLAST score 789 E value 4.0e-84 Match length 148 % identity 99

UBIQUITIN-CONJUGATING ENZYME E2-17 KD (UBIQUITIN-PROTEIN NCBI Description

LIGASE) (UBIQUITIN CARRIER PROTEIN) >gi 388207 (L23762)

ubiquitin carrier protein [Lycopersicon esculentum]

Seq. No. 5612

5134 2.R1040 Contig ID 5'-most EST pmv700890042.h1

Method BLASTX NCBI GI q464981 BLAST score 781 E value 2.0e-83 Match length 148 97 % identity

UBIQUITIN-CONJUGATING ENZYME E2-17 KD (UBIQUITIN-PROTEIN NCBI Description

LIGASE) (UBIQUITIN CARRIER PROTEIN) >qi 388207 (L23762)

ubiquitin carrier protein [Lycopersicon esculentum]

5613 Seq. No.

5134 3.R1040 Contig ID 5'-most EST zhf700961505.h1

BLASTN Method NCBI GI g456567 BLAST score 80 7.0e-37 E value Match length 200 % identity 85

NCBI Description Pisum sativum ubiquitin conjugating enzyme (UBC4), complete

cds

5614 Seq. No.

5134 4.R1040 Contig ID

5'-most EST uC-gmflminsoy056g03b1

Method BLASTX NCBI GI q441457 BLAST score 216 E value 2.0e-17 Match length 44 % identity

(X73419) ubiquitin conjugating enzyme E2 [Lycopersicon NCBI Description

esculentum]

5615 Seq. No.



Contig ID 5134 5.R1040 5'-most EST LIB3093-022-Q1-K1-A7 Method BLASTX NCBI GI 9464981

NCBI GI g464981 BLAST score 757 E value 2.0e-80 Match length 148 % identity 95

NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-17 KD (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN) >gi_388207 (L23762)

ubiquitin carrier protein [Lycopersicon esculentum]

Seq. No. 5616

Contig ID 5134_6.R1040 5'-most EST trc700561374.h1

Method BLASTN
NCBI GI g456567
BLAST score 48
E value 4.0e-18
Match length 80
% identity 90

NCBI Description Pisum sativum ubiquitin conjugating enzyme (UBC4), complete

cds

Seq. No. 5617

Contig ID 5135_2.R1040

Method BLASTN
NCBI GI g861154
BLAST score 130
E value 8.0e-67
Match length 294
% identity 86

NCBI Description V.faba VFCWINV1 mRNA for cell wall invertase I

Seq. No. 5618

Contig ID 5136_1.R1040 5'-most EST asn701139401.h1

Method BLASTX
NCBI GI g3023651
BLAST score 182
E value 1.0e-13
Match length 121

% identity 36

NCBI Description D-LACTATE DEHYDROGENASE [CYTOCHROME] PRECURSOR (D-LACTATE

FERRICYTOCHROME C OXIDOREDUCTASE) (D-LCR)

>gi 1076962_pir__S51528 D-lactate dehydrogenase

(cytochrome) (EC 1.1.2.4) - yeast (Kluyveromyces marxianus var. lactis) >gi 602029 emb CAA50635 (X71628) D-lactate

dehydrogenase (cytochrome) [Kluyveromyces lactis]

Seq. No. 5619

Contig ID 5138_1.R1040 5'-most EST wvk700682539.h2

Method BLASTX NCBI GI g3298441 BLAST score 405



```
E value
                   2.0e-39
Match length
                  140
                   62
% identity
NCBI Description
                  tabacum]
```

(AB010879) chloroplast ribosomal protein L10 [Nicotiana

Seq. No. Contig ID 5'-most EST 5620 5138 2.R1040 uxk700670986.h1

Method BLASTX g3298441 NCBI GI BLAST score 230 E value 2.0e-19 Match length 80 % identity 62

(AB010879) chloroplast ribosomal protein L10 [Nicotiana NCBI Description

tabacum]

Seq. No.

5621

Contig ID

5140 1.R1040

5'-most EST

LIB3051-078-Q1-K1-D12

Seq. No.

5622

Contig ID

5144 1.R1040

5'-most EST

LIB3028-005-Q1-B1-G7

Seq. No.

5623

Contig ID

5145 1.R1040 LIB3028-005-Q1-B1-G8

5'-most EST Method

BLASTX

NCBI GI BLAST score

q1076288 355

E value Match length 1.0e-33

% identity

104 62

NCBI Description amino acid permease AAP3 - Arabidopsis thaliana

Seq. No.

5624

Contig ID 5'-most EST 5149 1.R1040 ssr700559066.h1

Method NCBI GI BLASTX g4105794

BLAST score E value

442

Match length

2.0e-43 140

% identity

NCBI Description (AF049928) PGP224 [Petunia x hybrida]

Seq. No.

5625

57

Contig ID

5152 1.R1040

5'-most EST

LIB3028-005-Q1-B1-E11

Seq. No.

5626

Contig ID 5'-most EST

5152 2.R1040 wrg700787541.h1

Seq. No.

5627



Contig ID 5154_1.R1040

5'-most EST LIB3028-005-Q1-B1-E3

Seq. No. 5628

Contig ID 5159_1.R1040 5'-most EST asn701139905.h1

Method BLASTX
NCBI GI g3283996
BLAST score 1094
E value 1.0e-120
Match length 240
% identity 71

NCBI Description (AF072908) calcium-dependent protein kinase [Nicotiana

tabacum]

Seq. No. 5629

Contig ID 5162 1.R1040

5'-most EST uC-gmrominsoy310b03b1

Seq. No. 5630

Contig ID 5162 2.R1040

5'-most EST LIB3138-071-P1-N1-D2

Method BLASTX
NCBI GI g2708532
BLAST score 256
E value 8.0e-22
Match length 145
% identity 29

NCBI Description (AF029351) putative RNA binding protein [Nicotiana tabacum]

Seq. No. 5631

Contig ID 5163_1.R1040 5'-most EST zhf700955007.h1

Method BLASTX
NCBI GI g4567262
BLAST score 1083
E value 1.0e-118
Match length 258
% identity 79

NCBI Description (AC006841) putative ubiquitin [Arabidopsis thaliana]

Seq. No. 5632

Contig ID 5164_1.R1040

5'-most EST LIB3109-007-Q1-K1-D5

Method BLASTX
NCBI GI g2708737
BLAST score 718
E value 9.0e-76
Match length 295
% identity 40

NCBI Description (AC003952) putative nuclear protein [Arabidopsis thaliana]

Seq. No. 5633

Contig ID 5164_2.R1040

5'-most EST LIB3049-044-Q1-E1-B12

Method BLASTX NCBI GI g2708737



```
248
BLAST score
                  3.0e-21
E value
Match length
                  54
% identity
                  36
NCBI Description (AC003952) putative nuclear protein [Arabidopsis thaliana]
                  5634
Seq. No.
Contig ID
                  5166 1.R1040
                  jC-gmst02400009a02a1
5'-most EST
                  BLASTX
Method
                  g3080433
NCBI GI
                  558
BLAST score
                   3.0e-57
E value
Match length
                  196
% identity
                   50
                   (AL022605) putative gamma-glutamyltransferase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   5635
                   5171 1.R1040
Contig ID
                   zhf700962664.h1
5'-most EST
Method
                   BLASTX
                   g4455337
NCBI GI
                   155
BLAST score
E value
                   3.0e-10
                   100
Match length
                   40
% identity
                   (AL035525) pectinesterase-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   5636
Seq. No.
                   5172 1.R1040
Contig ID
                   gsv701052361.h1
5'-most EST
                   5637
Seq. No.
                   5173 1.R1040
Contig ID
5'-most EST
                   vwf700675859.h1
Seq. No.
                   5638
                   5175 1.R1040
Contig ID
                   xpa700796134.h1
5'-most EST
Method
                   BLASTN
NCBI GI
                   q557693
BLAST score
                   48
                   7.0e-18
E value
                   80
Match length
% identity
                   90
                   Arabidopsis thaliana Columbia GTP binding protein beta
NCBI Description
                   subunit (AGB1) mRNA, complete cds
                   5639
Seq. No.
                   5175 2.R1040
Contig ID
5'-most EST
                   pxt700944124.hl
```

Method BLASTX q1346106 NCBI GI 196 BLAST score

4.0e-15 E value



Match length % identity

NCBI Description

GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT >gi_557694 (U12232) GTP binding protein beta subunit [Arabidopsis thaliana] >gi_3096915_emb_CAA18825.1_ (AL023094) GTP binding protein beta subunit [Arabidopsis thaliana]

A STATE OF THE STA

Seq. No.

5176 1.R1040 Contig ID

uC-gmrominsoy103c05b1 5'-most EST

Seq. No.

5180 1.R1040 Contig ID hrw701059233.h1 5'-most EST

BLASTN Method g169295 NCBI GI 173 BLAST score 3.0e-92 E value Match length 471 84 % identity

Pharbitis nil heat shock protein 83 (Hsp83) gene, complete NCBI Description

cds

5642 Seq. No.

5181 1.R1040 Contig ID

LIB3139-110-P1-N1-E5 5'-most EST

BLASTX Method g951449 NCBI GI BLAST score 1060 1.0e-116 E value 324 Match length 59 % identity

(L46681) aspartic protease precursor [Lycopersicon NCBI Description

esculentum]

5643 Seq. No.

5184 1.R1040 Contig ID

uC-gmrominsoy128c08b1 5'-most EST

BLASTX Method q2244849 NCBI GI 417 BLAST score 1.0e-40 E value Match length 174 % identity

(Z97337) hypothetical protein [Arabidopsis thaliana] NCBI Description

5644 Seq. No.

5188 1.R1040 Contig ID 5'-most EST zhf700953137.h1

BLASTX Method g266567 NCBI GI BLAST score 845 E value 1.0e-90 Match length 265 % identity 65

MITOCHONDRIAL PROCESSING PEPTIDASE ALPHA SUBUNIT PRECURSOR NCBI Description



>gi_421956_pir__S23558 mitochondrial processing peptidase
(EC 3.4.99.41) alpha chain precursor - potato
>gi_21493_emb_CAA46990_ (X66284) mitochondrial processing
peptidase [Solanum tuberosum]

 Seq. No.
 5645

 Contig ID
 5188_2.R1040

 5'-most EST
 LIB3170-015-Q1-J1-C9

 Method
 BLASTX

 NCBI GI
 g266567

 BLAST score
 242

E value 3.0e-20
Match length 60
% identity 78

NCBI Description MITOCHONDRIAL PROCESSING PEPTIDASE ALPHA SUBUNIT PRECURSOR

(ALPHA-MPP) (UBIQUINOL-CYTOCHROME C REDUCTASE SUBUNIT II) >gi_421956_pir__S23558 mitochondrial processing peptidase

(EC 3.4.99.41) alpha chain precursor - potato

>gi_21493_emb_CAA46990_ (X66284) mitochondrial processing

peptidase [Solanum tuberosum]

Seq. No. 5646

Contig ID 5190_1.R1040

5'-most EST LIB3028-005-Q1-B1-B3

Seq. No. 5647

Contig ID 5191 1.R1040

5'-most EST uC-gmrominsoy131b04b1

Method BLASTX
NCBI GI 94128197
BLAST score 199
E value 2.0e-15
Match length 103
% identity 47

NCBI Description (U75273) acyl-CoA binding protein [Arabidopsis thaliana]

Seq. No. 5648

Contig ID 5191 2.R1040 5'-most EST cf1700863652.h1

Method BLASTX
NCBI GI g4128197
BLAST score 147
E value 3.0e-09
Match length 86
% identity 41

NCBI Description (U75273) acyl-CoA binding protein [Arabidopsis thaliana]

Seq. No. 5649

Contig ID 5192_1.R1040

5'-most EST uC-gmropic010d01b1

Method BLASTX
NCBI GI g602292
BLAST score 555
E value 5.0e-57
Match length 162
% identity 68

NCBI Description (U17987) RCH2 protein [Brassica napus]

E value

Match length

% identity

3.0e-23

112

50



```
5650
Seq. No.
                   5193 1.R1040
Contig ID
                   jC-gmf102220086f03a1
5'-most EST
                   BLASTX
Method
                   g1771162
NCBI GI
                   752
BLAST score
E value
                   5.0e-80
Match length
                   206
% identity
                  70
                   (X98930) SBT2 [Lycopersicon esculentum]
NCBI Description
                   >gi_3687307_emb_CAA07000_ (AJ006379) subtilisin-like
                   protease [Lycopersicon esculentum]
Seq. No.
                   5651
                   5200 1.R1040
Contig ID
                   vwf700675653.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2832672
                   623
BLAST score
E value
                   1.0e-64
Match length
                   148
                   82
% identity
                   (AL021712) nifU-like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5200 2.R1040
Contig ID
                   LIB3074-004-Q1-K1-F11
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2832672
                   629
BLAST score
                   2.0e-65
E value
                   153
Match length
                   82
% identity
                   (ALO21712) nifU-like protein [Arabidopsis thaliana]
NCBI Description
                   5653
Seq. No.
                   5202 1.R1040
Contig ID
                   fde700870929.h1
5'-most EST
                   BLASTX
Method
                   g3064045
NCBI GI
BLAST score
                   493
                   7.0e-50
E value
Match length
                   111
% identity
                   81
                   (AF054449) adenylosuccinate lyase [Mesembryanthemum
NCBI Description
                   crystallinum]
Seq. No.
                   5654
                   5213 1.R1040
Contig ID
5'-most EST
                   LIB3065-008-Q1-N1-H4
                   BLASTX
Method
NCBI GI
                   q1352681
BLAST score
                   269
```

1076



NCBI Description PROTEIN PHOSPHATASE 2C (PP2C) >gi_1076391_pir__S55457 phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis thaliana >gi_633028_dbj_BAA07287_ (D38109) protein phosphatase 2C [Arabidopsis thaliana]

• •

 Seq. No.
 5655

 Contig ID
 5213_2.R1040

 5'-most EST
 sat701007474.h2

Method BLASTX
NCBI GI g1352681
BLAST score 975
E value 1.0e-106
Match length 345
% identity 59

% identity 59
NCBI Description PROTEIN PHOSPHATASE 2C (PP2C) >gi_1076391_pir___S55457

phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis

thaliana >gi_633028_dbj_BAA07287_ (D38109) protein

phosphatase 2C [Arabidopsis thaliana]

Seq. No. 5656

Contig ID 5213_3.R1040

5'-most EST LIB3039-048-Q1-E1-A11

Method BLASTX
NCBI GI g1352681
BLAST score 196
E value 6.0e-15
Match length 128
% identity 43

NCBI Description PROTEIN PHOSPHATASE 2C (PP2C) >gi_1076391_pir__\$55457

phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis

thaliana >gi_633028_dbj_BAA07287_ (D38109) protein

phosphatase 2C [Arabidopsis thaliana]

Seq. No. 5657

Contig ID 5214_1.R1040 5'-most EST zhf700958201.h1

Method BLASTX
NCBI GI g1778015
BLAST score 209
E value 2.0e-16
Match length 99
% identity 43

NCBI Description (U59508) osmotic stress-induced proline dehydrogenase

[Arabidopsis thaliana]

Seq. No. 5658

Contig ID 5216_1.R1040 5'-most EST pmv700890843.h1

Method BLASTX
NCBI GI g1945611
BLAST score 920
E value 4.0e-99
Match length 435
% identity 42

NCBI Description (AB003103) 26S proteasome subunit p55 [Homo sapiens]

>gi_4506221_ref_NP_002807.1_pPSMD12_ proteasome (prosome,

macropain) 26S subunit, non-ATPase,



79

66

Match length % identity

NCBI Description

```
5659
Seq. No.
                  5216 2.R1040
Contig ID
                   jC-gmf102220051h08a1
5'-most EST
                  BLASTX
Method
                  g4432854
NCBI GI
BLAST score
                  239
                   3.0e-19
E value
                   208
Match length
                   18
% identity
                  (AC006300) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5660
Seq. No.
                   5216 3.R1040
Contig ID
                   zsg701119349.h1
5'-most EST
                   BLASTX
Method
                   g1945611
NCBI GI
BLAST score
                   169
                   1.0e-11
E value
                   66
Match length
                   48
% identity
                   (AB003103) 26S proteasome subunit p55 [Homo sapiens]
NCBI Description
                   >gi 4506221 ref NP 002807.1_pPSMD12_ proteasome (prosome,
                   macropain) 26S subunit, non-ATPase,
                   5661
Seq. No.
                   5220 1.R1040
Contig ID
                   uC-gmropic011h01b1
5'-most EST
                   BLASTX
Method
                   g3193298
NCBI GI
                   176
BLAST score
                   3.0e-12
E value
                   122
Match length
                   34
% identity
                   (AF069298) T14P8.17 gene product [Arabidopsis thaliana]
NCBI Description
                   5662
Seq. No.
                   5220 2.R1040
Contig ID
5'-most EST
                   LIB3040-041-Q1-E1-C5
                   BLASTX
Method
                   g3193298
NCBI GI
BLAST score
                   158
E value
                   2.0e-10
Match length
                   87
 % identity
                   (AF069298) T14P8.17 gene product [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5663
                   5226 1.R1040
 Contig ID
                   LIB3028-004-Q1-B1-G11
 5'-most EST
                   BLASTX
Method
                   q3287696
 NCBI GI
 BLAST score
                   257
 E value
                   3.0e-22
```

1078

(AC003979) Strong similarity to phosphoribosylanthranilate



transferase gb_D86180 from Pisum sativum. This ORF may be part of a larger gene that lies in the overlapping region. [Arabidopsis thaliana]

```
5664
Seq. No.
                  5229 1.R1040
Contig ID
                  hrw701058845.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3860321
BLAST score
                   1018
E value
                   1.0e-111
                   223
Match length
                   80
% identity
                   (AJ012687) beta-galactosidase [Cicer arietinum]
NCBI Description
Seq. No.
                   5233 1.R1040
Contig ID
                   leu701147642.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4469023
                   1492
BLAST score
                   1.0e-166
E value
Match length
                   387
% identity
                   76
                   (AL035602) putative protein [Arabidopsis thaliana]
NCBI Description
                   5666
Seq. No.
                   5233 2.R1040
Contig ID
                   jC-gmf102220051d01a1
5'-most EST
Method
                   BLASTX
                   q4469023
NCBI GI
                   744
BLAST score
                   6.0e-79
E value
                   199
Match length
                   73
% identity
                   (AL035602) putative protein [Arabidopsis thaliana]
NCBI Description
                   5667
Seq. No.
                   5235 1.R1040
Contig ID
                   LIB3107-057-Q1-K1-F4
5'-most EST
                   BLASTX
Method
                   g2739365
NCBI GI
                   801
BLAST score
                   2.0e-85
E value
Match length
                   188
                   72
% identity
                   (AC002505) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5668
Seq. No.
                   5235 2.R1040
Contig ID
 5'-most EST
                   g4276997
                   BLASTX
Method
                   q3201969
NCBI GI
BLAST score
                   299
E value
                   6.0e-27
```

1079

71

70

Match length

% identity



```
NCBI Description (AF068332) submergence induced protein 2A [Oryza sativa]
                  5669
Seq. No.
                  5235 4.R1040
Contig ID
                  LIB3028-047-Q1-B1-A10
5'-most EST
                  BLASTX
Method
                  g2739365
NCBI GI
BLAST score
                  423
E value
                   9.0e-42
                  101
Match length
% identity
                   72
                   (AC002505) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5670
Seq. No.
                   5239 1.R1040
Contig ID
                   g4396664
5'-most EST
                   BLASTX
Method
                   g3135264
NCBI GI
                   359
BLAST score
E value
                   3.0e-34
                   117
Match length
% identity
                  (AC003058) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5671
Seq. No.
                   5240 1.R1040
Contig ID
                   trc700563942.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g585338
BLAST score
                   1079
                   1.0e-118
E value
                   225
Match length
                   92
% identity
                   ADENYLATE KINASE B (ATP-AMP TRANSPHOSPHORYLASE)
NCBI Description
                   >gi_391879_dbj_BAA01181_ (D10335) adenylate kinase-b [Oryza
                   sativa]
                   5672
Seq. No.
                   5240 2.R1040
Contig ID
                   zzp700831593.h1
5'-most EST
                   BLASTX
Method
                   g585338
NCBI GI
                   419
BLAST score
                   3.0e-41
E value
                   89
Match length
                   91
 % identity
                   ADENYLATE KINASE B (ATP-AMP TRANSPHOSPHORYLASE)
NCBI Description
                   >gi 391879 dbj BAA01181_ (D10335) adenylate kinase-b [Oryza
                   sativa]
                   5673
 Seq. No.
                   5242 1.R1040
 Contig ID
```

5'-most EST $k117\overline{0}1211937.h1$

Seq. No. 5674

Contig ID 5243 1.R1040 pmv700893482.h1 5'-most EST



```
BLASTX
Method
NCBI GI
                  g3860251
                  1050
BLAST score
                  1.0e-115
E value
                  272
Match length
                  73
% identity
NCBI Description (AC005824) putative permease [Arabidopsis thaliana]
Seq. No.
                   5675
                   5244 1.R1040
Contig ID
                  uC-gmflminsoy029b03b1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g1280434
BLAST score
                   274
E value
                   1.0e-23
Match length
                   306
% identity
                   31
                  (U42014) hemomucin [Drosophila melanogaster]
NCBI Description
                   5676
Seq. No.
                   5245 1.R1040
Contig ID
                   leu701150315.h1
5'-most EST
                   5677
Seq. No.
                   5246 1.R1040
Contig ID
                   sat701009292.h1
5'-most EST
                   BLASTX
Method
                   q4455338
NCBI GI
BLAST score
                   910
E value
                   4.0e-98
                   272
Match length
                   64
% identity
                   (AL035525) putative protein [Arabidopsis thaliana]
NCBI Description
                   5678
Seq. No.
                   5250 1.R1040
Contig ID
5'-most EST
                   hyd700727832.h1
Seq. No.
                   5679
                   5255 1.R1040
Contig ID
                   epx701106273.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1176658
BLAST score
                   405
                   2.0e-60
E value
Match length
                   256
                   52
% identity
                   HYPOTHETICAL 200.6 KD PROTEIN B0228.2 IN CHROMOSOME II
NCBI Description
                   >gi 726363 (U23168) No definition line found
                   [Caenorhabditis elegans]
                   5680
Seq. No.
                   5255 2.R1040
Contig ID
```

Contig ID 5255_2.R1040 5'-most EST kmv700738802.h1

Method BLASTX NCBI GI g1176658

BLAST score 371



E value 6.0e-35
Match length 155
% identity 47

NCBI Description HYPOTHETICAL 200.6 KD PROTEIN B0228.2 IN CHROMOSOME II

>gi 726363 (U23168) No definition line found

[Caenorhabditis elegans]

Seq. No. 5681

Contig ID 5255_3.R1040

5'-most EST jC-gmf102220071a03a1

Seq. No. 5682

Contig ID 5255 4.R1040

5'-most EST LIB3051-034-Q1-K1-E12

Seq. No. 5683

Contig ID 5256_1.R1040

5'-most EST LIB3107-021-Q1-K1-G4

Seq. No. 5684

Contig ID 5256 2.R1040

5'-most EST LIB3092-030-Q1-K1-A12

Seq. No. 5685

Contig ID 5259 1.R1040

5'-most EST LIB3028-003-Q1-B1-H6

Method BLASTX
NCBI GI g3659909
BLAST score 281
E value 5.0e-25
Match length 80
% identity 69

NCBI Description (AF020715) histidyl-tRNA synthetase [Arabidopsis thaliana]

Seq. No. 5686

Contig ID 5261_1.R1040 5'-most EST rlr700898683.h1

Method BLASTX
NCBI GI g2832696
BLAST score 718
E value 6.0e-76
Match length 235
% identity 60

NCBI Description (AL021713) putative protein [Arabidopsis thaliana]

Seq. No. 5687

Contig ID 5262_1.R1040 5'-most EST zpv700759945.h1

Seq. No. 5688

Contig ID 5264_1.R1040 5'-most EST smc700748955.h1

Seq. No. 5689

Contig ID 5265 1.R1040

5'-most EST LIB3106-101-Q1-K1-G2

Method BLASTX



g2465923 NCBI GI 1068 BLAST score 1.0e-125 E value 461 Match length 54 % identity (AF024648) receptor-like serine/threonine kinase NCBI Description [Arabidopsis thaliana] 5690 Seq. No. 5265 3.R1040 Contig ID g5057646 5'-most EST 5691 Seq. No. 5266 1.R1040 Contig ID LIB3107-010-Q1-K1-H3 5'-most EST ${\tt BLASTX}$ Method q3292829 NCBI GI 1081 BLAST score 1.0e-118 E value 458 Match length 62 % identity (AL031018) putative protein [Arabidopsis thaliana] NCBI Description 5692 Seq. No. 5266 2.R1040 Contig ID LIB3030-002-Q1-B1-G9 5'-most EST Method BLASTX q2697129 NCBI GI BLAST score 426 4.0e-41 E value Match length 415 % identity (AF025654) mRNA capping enzyme [Homo sapiens] NCBI Description >gi_4506563_ref_NP_003791.1_pRNGTT_ UNKNOWN Seq. No. 5693 5269 1.R1040 Contig ID $leu7\overline{0}1156667.h1$ 5'-most EST Method BLASTX q2245020 NCBI GI BLAST score 184 4.0e-13 E value 195 Match length 30 % identity (Z97341) growth regulator homolog [Arabidopsis thaliana] NCBI Description

5694 Seq. No.

5269 2.R1040 Contig ID uaw700666916.h1 5'-most EST

Seq. No. 5695

5271 1.R1040 Contig ID $dpv7\overline{0}1096923.h1$ 5'-most EST

Method BLASTX NCBI GI g3820614 BLAST score 480 4.0e-48 E value



189 Match length 55 % identity

(AF094516) El-like protein [Homo sapiens] NCBI Description

5696 Seq. No.

5272 1.R1040 Contig ID

5'-most EST jC-gmro02910016d05a1

BLASTX Method NCBI GI q4539371 BLAST score 410 1.0e-39 E value 199 Match length

50 % identity

(AL049525) putative protein [Arabidopsis thaliana] NCBI Description

5697 Seq. No.

5272 2.R1040 Contig ID

jC-gmle01810005e04a1 5'-most EST

Seq. No. 5698

5276 1.R1040 Contig ID

jC-gmle01810082c09a1 5'-most EST

BLASTX Method g3043529 NCBI GI 292 BLAST score 3.0e-26 E value 92 Match length % identity

NCBI Description (AJ002204) polyamine oxidase [Zea mays]

5699 Seq. No.

5277 1.R1040 Contig ID hrw701063235.h1 5'-most EST

BLASTX Method q4063751 NCBI GI 145 BLAST score 2.0e-13 E value 118 Match length 47 % identity

(AC005851) putative white protein [Arabidopsis thaliana] NCBI Description >gi_4510409_gb_AAD21495.1_ (AC006929) putative white

protein [Arabidopsis thaliana]

Seq. No. 5700

5279 1.R1040 Contig ID

uC-gmflminsoy079g08b1 5'-most EST

BLASTX Method NCBI GI g2288997 BLAST score 286 1.0e-25 E value 69 Match length % identity 80

(AC002335) hypothetical protein [Arabidopsis thaliana] NCBI Description

5701 Seq. No.

5280 1.R1040 Contig ID

LIB3049-002-Q1-E1-F7 5'-most EST



```
BLASTX
Method
                  g2632252
NCBI GI
                  558
BLAST score
                  4.0e-57
E value
                  186
Match length
% identity
                   61
NCBI Description (Y12464) serine/threonine kinase [Sorghum bicolor]
Seq. No.
                   5702
                   5280 2.R1040
Contig ID
                   jC-qmst02400056c02a1
5'-most EST
                   BLASTX
Method
                   q2632252
NCBI GI
BLAST score
                   586
                   2.0e-60
E value
                   195
Match length
                   61
% identity
                   (Y12464) serine/threonine kinase [Sorghum bicolor]
NCBI Description
                   5703
Seq. No.
                   5284 1.R1040
Contig ID
                   jC-gmro02910060c10d1
5'-most EST
                   5704
Seq. No.
                   5289 1.R1040
Contig ID
                   LIB3094-023-Q1-K1-B2
5'-most EST
                   BLASTX
Method
                   g3142331
NCBI GI
                   237
BLAST score
                   2.0e-19
E value
                   178
Match length
                   33
% identity
                  (U97327) calcyclin binding protein [Mus musculus]
NCBI Description
Seq. No.
                   5705
                   5289 2.R1040
Contig ID
                   kl1701206226.h1
5'-most EST
                   BLASTX
Method
                   q3122258
NCBI GI
```

BLAST score 916

6.0e-99 E value Match length 245 % identity 73

EUKARYOTIC TRANSLATION INITIATION FACTOR 6 (EIF-6) (B4 NCBI Description INTEGRIN INTERACTOR) (CAB) >gi 2809383 (AF022229)

translation initiation factor 6 [Homo sapiens]

>gi_2910997_emb_CAA72243_ (Y11435) b4 integrin interactor [Homo sapiens] >gi 3335506 (AF047433) b(2)gcn homolog [Homo sapiens] >gi_4504771_ref_NP_002203.1_pITGB4BP_ integrin

beta 4 binding protein

5706 Seq. No.

Contig ID 5289 3.R1040 5'-most EST zhf700956194.h1

BLASTX Method q3355311 NCBI GI 564 BLAST score



E value 5.0e-58
Match length 128
% identity 85

NCBI Description (AJ009737) eukaryotic translation initiation factor 6 [Beta

vulgaris]

Seq. No. 5707

Contig ID 5289 5.R1040

5'-most EST LIB3040-055-Q1-E1-H11

Method BLASTX
NCBI GI g3355311
BLAST score 278
E value 7.0e-25
Match length 59
% identity 92

NCBI Description (AJ009737) eukaryotic translation initiation factor 6 [Beta

vulgaris]

Seq. No. 5708

Contig ID 5289_10.R1040 5'-most EST smc700750180.h1

Method BLASTN
NCBI GI g3355310
BLAST score 45
E value 3.0e-16
Match length 93
% identity 87

NCBI Description Beta vulgaris mRNA for eukaryotic translation initiation

factor 6, partial

Seq. No. 5709

Contig ID 5290_1.R1040

5'-most EST LIB3028-003-Q1-B1-E3

Method BLASTX
NCBI GI g2252830
BLAST score 829
E value 1.0e-88
Match length 289
% identity 57

NCBI Description (AF013293) weak similarity to receptor protein kinase

[Arabidopsis thaliana]

Seq. No. 5710

Contig ID 5301 1.R1040

5'-most EST LIB3170-013-Q1-K1-D5

Method BLASTX
NCBI GI g4006872
BLAST score 994
E value 1.0e-108
Match length 283
% identity 65

NCBI Description (Z99707) methionyl aminopeptidase-like protein [Arabidopsis]

thaliana]

Seq. No. 5711

Contig ID 5302 1.R1040 5'-most EST wrg700789428.h2



```
BLASTX
Method
                  g2829910
NCBI GI
                  523
BLAST score
                  1.0e-53
E value
Match length
                  114
% identity
                  21
                   (AC002291) Unknown protein, contains regulator of
NCBI Description
                  chromosome condensation motifs [Arabidopsis thaliana]
Seq. No.
                  5712
                  5304 1.R1040
Contig ID
                  pxt700944576.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2129496
BLAST score
                  1529
                  0.0e+00
E value
Match length
                   424
% identity
                   81
                   acetolactate synthase (EC 4.1.3.18) precursor (clone A19) -
NCBI Description
                   upland cotton
Seq. No.
                   5713
                   5305 1.R1040
Contig ID
                   LIB3028-003-Q1-B1-F8
5'-most EST
Seq. No.
                   5714
                   5308 1.R1040
Contig ID
                   zhf700954109.hl
5'-most EST
                   BLASTN
Method
                   g3617740
NCBI GI
                   37
BLAST score
                   7.0e-11
E value
                   157
Match length
                   81
% identity
                   Arabidopsis thaliana BAC F1I21 from chromosome 1, near 59
NCBI Description
                   cM, complete sequence [Arabidopsis thaliana]
                   5715
Seq. No.
                   5308 2.R1040
Contig ID
                   LIB3051-088-Q1-K1-C3
5'-most EST
Method
                   BLASTN
NCBI GI
                   g2281636
BLAST score
                   37
                   3.0e-11
E value
                   157
Match length
% identity
                   Arabidopsis thaliana AP2 domain containing protein RAP2.6
NCBI Description
                   mRNA, partial cds
                   5716
Seq. No.
                   5308 3.R1040
Contig ID
                   jC-gmf102220089g10a1
5'-most EST
```

Seq. No. 5717

Contig ID 5'-most EST

5313_1.R1040 jex700904162.h1

Method BLASTX



NCBI GI g2149640 BLAST score 1365 E value 1.0e-151 Match length 316 % identity 84

NCBI Description (U91995) Argonaute protein [Arabidopsis thaliana]

Seq. No. 5718

Contig ID 5313_2.R1040

5'-most EST uC-qmflminsoy026c06b1

Method BLASTX
NCBI GI g2149640
BLAST score 931
E value 1.0e-101
Match length 206
% identity 86

NCBI Description (U91995) Argonaute protein [Arabidopsis thaliana]

Seq. No. 5719

Contig ID 5313_3.R1040

5'-most EST LIB3107-014-Q1-K1-F3

Method BLASTX
NCBI GI g2149640
BLAST score 1754
E value 0.0e+00
Match length 393
% identity 86

NCBI Description (U91995) Argonaute protein [Arabidopsis thaliana]

Seq. No. 5720

Contig ID 5313_4.R1040 5'-most EST jsh701069173.h1

Method BLASTX
NCBI GI g2149640
BLAST score 175
E value 3.0e-22
Match length 75
% identity 77

NCBI Description (U91995) Argonaute protein [Arabidopsis thaliana]

Seq. No. 5721

Contig ID 5315_1.R1040 5'-most EST wrg700786111.h2

Method BLASTX
NCBI GI g1710401
BLAST score 594
E value 1.0e-90
Match length 189
% identity 87

NCBI Description RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE SMALL CHAIN

(RIBONUCLEOTIDE REDUCTASE) (R2 SUBUNIT)

>gi_1044912_emb_CAA63194_ (X92443) ribonucleotide reductase

R2 [Nicotiana tabacum]

Seq. No. 5722

Contig ID 5317 1.R1040

5'-most EST LIB3028-003-Q1-B1-D8



Method BLASTX
NCBI GI 9466158
BLAST score 219
E value 2.0e-23
Match length 74
% identity 70

NCBI Description HYPOTHETICAL 9.3 KD PROTEIN ZK652.1 IN CHROMOSOME III >gi_630768_pir__S44901 ZK652.1 protein - Caenorhabditis elegans >gi_289767 (L14429) putative [Caenorhabditis

elegans]

Seq. No. 5723

Contig ID 5318_1.R1040 5'-most EST ncj700987601.h1

Method BLASTN
NCBI GI g3510539
BLAST score 291
E value 1.0e-162
Match length 591
% identity 87

NCBI Description Prunus armeniaca expansin (Exp2) mRNA, complete cds

Seq. No. 5724

Contig ID 5319_1.R1040 5'-most EST bth700849223.h1

Seq. No. 5725

Contig ID 5320 1.R1040

5'-most EST LIB3028-003-Q1-B1-E10

Seq. No. 5726

Contig ID 5324_1.R1040 5'-most EST zsg701122415.h1

Seq. No. 5727

Contig ID 5326_1.R1040 5'-most EST ncj700983837.h1

Method BLASTX
NCBI GI g2960216
BLAST score 1563
E value 1.0e-174
Match length 382
% identity 77

NCBI Description (AJ223384) 26S proteasome regulatory ATPase subunit 10b

(S10b) [Manduca sexta]

Seq. No. 5728

Contig ID 5326_2.R1040 5'-most EST jsh701064850.h1

Method BLASTN
NCBI GI g3236477
BLAST score 123
E value 1.0e-62
Match length 227
% identity 89

NCBI Description Gossypium hirsutum 26S protease regulatory subunit (PRA2)

mRNA, partial cds



```
5729
Seq. No.
                  5329 2.R1040
Contig ID
                  smc700744869.hl
5'-most EST
                  5730
Seq. No.
Contig ID
                  5330 1.R1040
5'-most EST
                  LIB3028-003-Q1-B1-B4
Seq. No.
                  5332 1.R1040
Contig ID
                  gsv701054195.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4454469
BLAST score
                   636
                   1.0e-66
E value
Match length
                   151
                   75
% identity
NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]
                   5732
Seq. No.
                   5338 1.R1040
Contig ID
5'-most EST
                   LIB3073-003-Q1-K1-A7
Method
                   BLASTX
NCBI GI
                   q3608481
BLAST score
                   621
E value
                   1.0e-64
Match length
                   147
% identity
                   80
                   (AF088913) ribosomal protein L27a [Petunia x hybrida]
NCBI Description
                   5733
Seq. No.
                   5338 2.R1040
Contig ID
                   LIB3049-034-Q1-E1-H7
5'-most EST
Method
                   BLASTX
                   q3608481
NCBI GI
BLAST score
                   590
                   6.0e-61
E value
                   142
Match length
                   78
% identity
                   (AF088913) ribosomal protein L27a [Petunia x hybrida]
NCBI Description
                   5734
Seq. No.
                   5348 1.R1040
Contig ID
                   qsv701055090.h1
5'-most EST
                   BLASTX
Method
                   g2583117
NCBI GI
                   220
BLAST score
                   2.0e-17
E value
Match length
                   193
% identity
                   (AC002387) hypothetical protein [Arabidopsis thaliana]
NCBI Description
```

Seq. No.

Contig ID 5348 2.R1040 5'-most EST

uC-qmropic100f03b1

Method BLASTX

NCBI Description

Seq. No.

5743



```
q2827551
NCBI GI
                  202
BLAST score
                  2.0e-15
E value
                  133
Match length
                  42
% identity
                  (AL021635) predicted protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  5350 1.R1040
Contig ID
                  jC-gmro02800032c07a1
5'-most EST
                  BLASTX
Method
                  q3445201
NCBI GI
BLAST score
                  246
                   2.0e-20
E value
                  154
Match length
                   40
% identity
NCBI Description (AC004786) unknown protein [Arabidopsis thaliana]
Seq. No.
                   5737
                   5351 1.R1040
Contig ID
                   g5677092
5'-most EST
                   5738
Seq. No.
                   5351 2.R1040
Contig ID
5'-most EST
                   pcp700992609.hl
                   5739
Seq. No.
                   5351 3.R1040
Contig ID
5'-most EST
                   hyd700725439.h1
                   5740
Seq. No.
                   5357 1.R1040
Contig ID
                   uC-gmrominsoy194a10b1
5'-most EST
                   5741
Seq. No.
                   5359 1.R1040
Contig ID
                   LIB3028-002-Q1-B1-F7
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3269292
                   280
BLAST score
                   1.0e-24
E value
Match length
                   72
% identity
                   (AL030978) putative protein [Arabidopsis thaliana]
NCBI Description
                   5742
Seq. No.
                   5362 1.R1040
Contig ID
5'-most EST
                   LIB3028-002-Q1-B1-G1
Method
                   BLASTX
NCBI GI
                   g2979556
BLAST score
                   127
E value
                   9.0e-12
Match length
                   157
% identity
                   33
```

(AC003680) unknown protein [Arabidopsis thaliana]



```
5363 1.R1040
Contig ID
                  txt700732041.h1
5'-most EST
                  BLASTX
Method
                  g2352492
NCBI GI
                  310
BLAST score
                  4.0e-28
E value
                  127
Match length
                   46
% identity
                   (AF005047) transport inhibitor response 1 [Arabidopsis
NCBI Description
                   thaliana] >gi 2352494 (AF005048) transport inhibitor
                   response 1 [Arabidopsis thaliana]
                   5744
Seq. No.
                   5363 2.R1040
Contig ID
                   dpv701097370.hl
5'-most EST
                   BLASTX
Method
                   g2352492
NCBI GI
                   309
BLAST score
E value
                   3.0e-28
Match length
                   123
                   50
% identity
                   (AF005047) transport inhibitor response 1 [Arabidopsis
NCBI Description
                   thaliana] >gi 2352494 (AF005048) transport inhibitor
                   response 1 [Arabidopsis thaliana]
                   5745
Seq. No.
                   5363 3.R1040
Contig ID
                   gsv701046587.h1
5'-most EST
                   BLASTX
Method
                   g4262147
NCBI GI
                   148
BLAST score
                   9.0e-10
E value
Match length
                   44
                   59
% identity
                   (AC005275) putative homolog of transport inhibitor response
NCBI Description
                   1 [Arabidopsis thaliana]
                   5746
Seq. No.
                   5364 1.R1040
Contig ID
                   g5508964
5'-most EST
                   BLASTN
Method
NCBI GI
                   g21103
BLAST score
                   72
E value
                   4.0e-32
                   180
Match length
                   85
 % identity
                   R.sativus mRNA for HMG-COA reductase (clone cRS8)
NCBI Description
Seq. No.
                   5747
                   5364 2.R1040
Contig ID
                   zhf700954225.h1
 5'-most EST
```

NCBI GI g21103 BLAST score 56 E value 1.0e-22 Match length 180 % identity 83

Method

BLASTN



```
NCBI Description R.sativus mRNA for HMG-COA reductase (clone cRS8)
                  5748
Seq. No.
                  5365 1.R1040
Contig ID
                  LIB3028-002-Q1-B1-G12
5'-most EST
                  5749
Seq. No.
                  5367 1.R1040
Contig ID
5'-most EST
                  LIB3028-002-Q1-B1-B6
                  BLASTX
Method
NCBI GI
                   q3914006
BLAST score
                   552
                   2.0e-67
E value
Match length
                   242
% identity
                   61
                  MITOCHONDRIAL LON PROTEASE HOMOLOG 2 PRECURSOR >gi 1816588
NCBI Description
                   (U85495) LON2 [Zea mays]
                   5750
Seq. No.
                   5372 1.R1040
Contig ID
                   LIB3094-054-Q1-K1-B6
5'-most EST
                   BLASTX
Method
                   q3023728
NCBI GI
BLAST score
                   290
                   2.0e-25
E value
Match length
                   246
                   31
% identity
                   EXOSTOSIN-2 (PUTATIVE TUMOUR SUPPRESSOR PROTEIN EXT2)
NCBI Description
                   (MULTIPLE EXOSTOSES PROTEIN 2) >gi_1619954 (U72141)
                   multiple exostosis 2 protein [Mus musculus]
                   5751
Seq. No.
                   5373 1.R1040
Contig ID
                   kl1701214081.h1
5'-most EST
                   BLASTX
Method
                   q4539335
NCBI GI
                   1454
BLAST score
                   1.0e-176
E value
                   486
Match length
                   60
% identity
NCBI Description (AL035539) putative protein [Arabidopsis thaliana]
                   5752
Seq. No.
Contig ID
                   5374 1.R1040
 5'-most EST
                   LIB3109-034-Q1-K1-B1
                   5753
 Seq. No.
                   5375 1.R1040
Contig ID
 5'-most EST
                   LIB3028-002-Q1-B1-D9
                   BLASTX
Method
                   q2499570
 NCBI GI
                   394
 BLAST score
                   4.0e-49
 E value
 Match length
                   147
 % identity
                    68
                   PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE
 NCBI Description
```

(PROTEIN-BETA-ASPARTATE METHYLTRANSFERASE) (PIMT) (PROTEIN



L-ISOASPARTYL METHYLTRANSFERASE) (L-ISOASPARTYL PROTEIN CARBOXYL METHYLTRANSFERASE) >gi_414332 (L07941) L-isoaspartyl methyltransferase [Triticum aestivum]

5754 Seq. No. 5377 1.R1040 Contig ID leu701147682.hl 5'-most EST BLASTX Method NCBI GI g2879867 BLAST score 415 E value 3.0e-40Match length 134 62 % identity (AL021816) 40s ribosomal protein S17 [Schizosaccharomyces NCBI Description pombe] 5755 Seq. No. 5377 2.R1040 Contig ID 5'-most EST LIB3170-005-Q1-K1-E3 Method BLASTX NCBI GI g2879867 BLAST score 420 E value 6.0e-41Match length 134 % identity 63 (AL021816) 40s ribosomal protein S17 [Schizosaccharomyces NCBI Description pombe] 5756 Seq. No. 5377 3.R1040 Contig ID LIB3040-026-Q1-E1-D1 5'-most EST BLASTN Method NCBI GI g2618605 BLAST score 110 E value 1.0e-54 262 Match length 85 % identity Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description MUK11, complete sequence [Arabidopsis thaliana] 5757 Seq. No. 5377 4.R1040 Contig ID LIB3049-054-Q1-E1-G3 5'-most EST BLASTN Method NCBI GI g2618605 BLAST score 89 2.0e-42 E value 305 Match length 82 % identity Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description

Seq. No. Contig ID

5758 5380_1.R1040 leu701148742.h1

5'-most EST leu7011 Method BLASTX NCBI GI q228403

1094

MUK11, complete sequence [Arabidopsis thaliana]



BLAST score 1676 E value 0.0e+00 Match length 371 % identity 89

NCBI Description glycolate oxidase [Lens culinaris]

Seq. No. 5759

Contig ID 5380 2.R1040

5'-most EST LIB3093-032-Q1-K1-G8

Method BLASTN
NCBI GI g3435305
BLAST score 94
E value 4.0e-45
Match length 266
% identity 84

NCBI Description Medicago sativa glycolate oxidase mRNA, partial cds

Seq. No. 5760

Contig ID 5382 1.R1040

5'-most EST LIB3053-005-Q1-N1-G12

Method BLASTN
NCBI GI g508303
BLAST score 83
E value 2.0e-38
Match length 243
% identity 84

NCBI Description Alfalfa bimodular protein (corC) mRNA, complete cds

Seq. No. 5761

Contig ID 5382 2.R1040

5'-most EST LIB3\overline{106}-043-\overline{Q1}-K1-D1

Method BLASTN
NCBI GI g508303
BLAST score 81
E value 1.0e-37
Match length 185
% identity 86

NCBI Description Alfalfa bimodular protein (corC) mRNA, complete cds

Seq. No. 5762

Contig ID 5383 1.R1040

5'-most EST LIB3049-051-Q1-E1-D6

Seq. No. 5763

Contig ID 5384 1.R1040

5'-most EST LIB3028-009-Q1-B1-E3

Method BLASTN
NCBI GI g563328
BLAST score 120
E value 1.0e-60
Match length 316
% identity 86

NCBI Description A.officinalis mRNA for histone 2B

Seq. No. 5764

Contig ID 5384_2.R1040

5'-most EST LIB3040-058-Q1-E1-C11



Method BLASTN
NCBI GI g2935449
BLAST score 114
E value 4.0e-57
Match length 278
% identity 85

NCBI Description Malus domestica histone H2B mRNA, partial cds

Seq. No. 5765

Contig ID 5384 3.R1040

5'-most EST LIB3170-001-Q1-K1-D10

Method BLASTN
NCBI GI g563328
BLAST score 138
E value 2.0e-71
Match length 310
% identity 88

NCBI Description A.officinalis mRNA for histone 2B

Seq. No. 5766

Contig ID 5384 4.R1040

5'-most EST LIB3040-032-Q1-E1-F8

Method BLASTN
NCBI GI g563328
BLAST score 136
E value 4.0e-70
Match length 280
% identity 87

NCBI Description A.officinalis mRNA for histone 2B

Seq. No. 5767

Contig ID 5384 5.R1040

5'-most EST LIB3106-014-Q1-K1-D4

Method BLASTN
NCBI GI g2558961
BLAST score 129
E value 6.0e-66
Match length 265
% identity 87

NCBI Description Gossypium hirsutum histone H2B1 mRNA, complete cds

Seq. No. 5768

Contig ID 5384 6.R1040

5'-most EST jC-gmro02910069a02a1

Method BLASTN
NCBI GI g563328
BLAST score 127
E value 8.0e-65
Match length 315
% identity 87

NCBI Description A.officinalis mRNA for histone 2B

Seq. No. 5769

Contig ID 5384_7.R1040

5'-most EST LIB3039-012-Q1-E1-E7

Method BLASTN NCBI GI g2935449



```
BLAST score
                   123
E value
                   2.0e-62
Match length
                   279
                   86
% identity
```

NCBI Description Malus domestica histone H2B mRNA, partial cds

Seq. No.

5770

Contig ID 5'-most EST 5384 8.R1040

Method

LIB3039-047-Q1-E1-H1

BLASTN NCBI GI q2558961 BLAST score 120 1.0e-60 E value Match length 276 86 % identity

NCBI Description Gossypium hirsutum histone H2B1 mRNA, complete cds

Seq. No.

5771

Contig ID

5384 10.R1040

5'-most EST

LIB3049-040-Q1-E1-F7

Method BLASTN NCBI GI q2935449 BLAST score 82 E value 3.0e-38 Match length 114 % identity 93

NCBI Description Malus domestica histone H2B mRNA, partial cds

Seq. No.

5772

Contig ID

5384 11.R1040

5'-most EST

LIB3106-033-Q1-K1-F3

Method BLASTN NCBI GI g2935449 BLAST score 88 E value 8.0e-42 Match length 160 89

% identity

NCBI Description Malus domestica histone H2B mRNA, partial cds

Seq. No.

5773

Contig ID

5384 12.R1040

LIB3170-002-Q1-K1-G8

5'-most EST Method

BLASTN

NCBI GI BLAST score g4309679 75

E value

5.0e-34

Match length

207

% identity

43

NCBI Description Arabidopsis thaliana phosphoribosylanthranilate isomerase

(PAI1) and (PAI4) genes, complete cds

Seq. No.

5774

Contig ID 5'-most EST

5385 1.R1040 vzy700755223.h1

Method BLASTX g2529707 NCBI GI BLAST score 430



```
E value
                  2.0e-42
Match length
                  148
                  55
% identity
                  (AF001434) Hpast [Homo sapiens]
NCBI Description
                  5775
Seq. No.
Contig ID
                  5387 1.R1040
5'-most EST
                  vwf700677464.hl
Method
                  BLASTX
NCBI GI
                  q1762436
BLAST score
                  1379
E value
                  1.0e-153
Match length
                  353
% identity
                  72
                   (U60277) acyl-acyl carrier protein desaturase [Asclepias
NCBI Description
                  syriaca]
Seq. No.
                  5776
                  5390 1.R1040
Contig ID
5'-most EST
                  LIB3138-017-Q1-N1-C12
Method
                  BLASTX
NCBI GI
                  q3451078
BLAST score
                  306
E value
                  2.0e-27
Match length
                  282
% identity
                  33
NCBI Description
                  (AL031326) putative protein [Arabidopsis thaliana]
Seq. No.
                  5777
                  5390 2.R1040
Contig ID
5'-most EST
                  LIB3051-102-Q1-K1-H1
                  5778
Seq. No.
                  5395 1.R1040
Contig ID
5'-most EST
                  sat701014384.h1
Method
                  BLASTX
NCBI GI
                  q4204281
BLAST score
                  338
E value
                  2.0e-31
Match length
                  175
% identity
NCBI Description
                  (AC004146) Hypothetical protein [Arabidopsis thaliana]
                  5779
Seq. No.
                  5395 2.R1040
Contig ID
5'-most EST
                  LIB3051-091-Q1-K1-G11
Seq. No.
                  5780
                  5412 1.R1040
Contig ID
5'-most EST
                  LIB3094-090-Q1-K1-D2
Method
                  BLASTX
```

Method BLASTX
NCBI GI g2827551
BLAST score 456
E value 2.0e-45
Match length 115
% identity 66

NCBI Description (AL021635) predicted protein [Arabidopsis thaliana]



Seq. No. 5781

5414 1.R1040 Contig ID 5'-most EST crh700854530.h1

Method BLASTX NCBI GI g1944518 BLAST score 1142 1.0e-125 E value Match length 303 % identity

NCBI Description (Y07822) Shaqqy-like kinase tetha [Arabidopsis thaliana]

>qi 3047105 (AF058919) protein kinase [Arabidopsis

thaliana]

Seg. No.

5419 1.R1040 Contig ID

5'-most EST LIB3039-039-Q1-E1-D5

5782

Method BLASTX NCBI GI q3201623 BLAST score 1864 E value 0.0e+00Match length 380 % identity 91

NCBI Description (AC004669) shaggy-like kinase dzeta [Arabidopsis thaliana]

5783 Seq. No.

Contig ID 5419 2.R1040

5'-most EST LIB3093-036-Q1-K1-E5

Method BLASTX NCBI GI q1877397 BLAST score 306 1.0e-27 E value 102 Match length 58 % identity

(Y11591) shaggy-like kinase [Ricinus communis] NCBI Description

Seq. No.

5784

Contig ID 5419 4.R1040

5'-most EST LIB3170-006-Q1-K1-B6

Method BLASTX NCBI GI q2129738 BLAST score 484 E value 1.0e-48 Match length 117 % identity 81

NCBI Description

shaggy-like kinase dzeta - Arabidopsis thaliana >gi_1225913_emb_CAA64408_ (X94938) shaggy-like kinase dzeta [Arabidopsis thaliana] >gi_1669653_emb_CAA70483_ (Y09300)

serine/threonine kinase [Arabidopsis thaliana]

5785 Seq. No.

Contig ID 5419 5.R1040

5'-most EST LIB3109-041-Q1-K1-G7

Method BLASTN NCBI GI g1877396 BLAST score 106 E value 1.0e-52



Match length 274 % identity 85

NCBI Description R.communis mRNA for shaggy-like kinase, partial

Seq. No. 5786

Contig ID 5419_6.R1040 5'-most EST zhf700953024.h1

Method BLASTN
NCBI GI g1161509
BLAST score 56
E value 1.0e-22
Match length 140
% identity 85

NCBI Description A.thaliana mRNA for shaggy-like kinase dzeta

Seq. No. 5787

Contig ID 5425_1.R1040

5'-most EST LIB3028-001-Q1-B1-H3

Seq. No. 5788

Contig ID 5427_1.R1040 5'-most EST fua701038054.h1

Method BLASTX
NCBI GI g2982311
BLAST score 867
E value 4.0e-93
Match length 238
% identity 68

NCBI Description (AF051240) probable ubiquitin-conjugating enzyme E2 [Picea

mariana]

Seq. No. 5789

Contig ID 5429 1.R1040

5'-most EST LIB3056-004-Q1-N1-H3

Seq. No. 5790

Contig ID 5431_1.R1040 5'-most EST ncj700975521.h1

Method BLASTX
NCBI GI g2160166
BLAST score 631
E value 2.0e-65
Match length 243
% identity 51

NCBI Description (AC000132) No definition line found [Arabidopsis thaliana]

Seq. No. 5791

Contig ID 5434_1.R1040

5'-most EST LIB3049-003-Q1-E1-F11

Method BLASTX
NCBI GI g3850582
BLAST score 281
E value 9.0e-25
Match length 64
% identity 81

NCBI Description (AC005278) F15K9.14 [Arabidopsis thaliana]

Seq. No.

Contig ID

5'-most EST

5798

5445 1.R1040

LIB3051-101-Q1-K1-F3



```
5792
Seq. No.
                  5434 2.R1040
Contig ID
5'-most EST
                  sat701015329.hl
                  BLASTX
Method
NCBI GI
                  a3850582
BLAST score
                  249
E value
                  4.0e-21
Match length
                  55
% identity
                  84
NCBI Description (AC005278) F15K9.14 [Arabidopsis thaliana]
                  5793
Seq. No.
Contig ID
                  5435 1.R1040
5'-most EST
                  wvk700685378.hl
Method
                  BLASTX
NCBI GI
                  q1706476
BLAST score
                  199
E value
                  3.0e-15
Match length
                  71
% identity
                  55
NCBI Description DNAJ-LIKE PROTEIN SLR0093 >gi 1001729 dbj BAA10566_
                  (D64004) DnaJ [Synechocystis sp.]
Seq. No.
                  5794
Contig ID
                  5438 1.R1040
5'-most EST
                  LIB3028-001-Q1-B1-F7
Method
                  BLASTX
NCBI GI
                  g3892054
BLAST score
                  536
E value
                  6.0e-55
Match length
                  133
% identity
                  67
NCBI Description (AC002330) putative glycosyltransferase [Arabidopsis
                  thaliana]
Seq. No.
                  5795
                  5439 1.R1040
Contig ID
                  LIB3028-001-Q1-B1-G1
5'-most EST
Seq. No.
                  5796
Contig ID
                  5442 1.R1040
5'-most EST
                  uxk700671601.h1
Method
                  BLASTX
NCBI GI
                  q4538964
BLAST score
                  148
E value
                  3.0e-09
Match length
                  58
% identity
                  47
NCBI Description (AL049488) NAM/NAP like protein [Arabidopsis thaliana]
                  5797
Seq. No.
                  5442 2.R1040
Contig ID
5'-most EST
                  zhf700956093.h1
```



Seq. No. 5799

Contig ID 5445 2.R1040

5'-most EST LIB3050-012-Q1-E1-G5

Seq. No. 5800

Contig ID 5445_3.R1040 5'-most EST pxt700942928.h1

Seq. No. 5801

Contig ID 5445_4.R1040 5'-most EST uxk700668613.h1

Seq. No. 5802

Contig ID 5447 1.R1040

5'-most EST LIB3028-001-Q1-B1-D8

Seq. No. 5803

Contig ID 5448_1.R1040

5'-most EST fC-gmse7000756618r1

Method BLASTX
NCBI GI g4006924
BLAST score 267
E value 4.0e-23
Match length 60
% identity 75

NCBI Description (Z99708) beta-galactosidase like protein [Arabidopsis

thaliana]

Seq. No. 5804

Contig ID 5448_2.R1040

5'-most EST fC-gmse700756174a2

Method BLASTX
NCBI GI g2580440
BLAST score 293
E value 4.0e-26
Match length 81
% identity 72

NCBI Description (D87261) PCF2 [Oryza sativa]

Seq. No. 5805

Contig ID 5448 3.R1040

5'-most EST uC-gmropic030d07b1

Method BLASTN
NCBI GI g2580437
BLAST score 68
E value 9.0e-30
Match length 164
% identity 85

NCBI Description Oryza sativa mRNA for PCF1, complete cds

Seq. No. 5806

Contig ID 5448_4.R1040 5'-most EST g5753241

Seq. No. 5807

Contig ID 5448_5.R1040



5'-most EST fC-qmse7000756174f1 Method BLASTX NCBI GI g2979559 145 BLAST score 9.0e-21 E value Match length 171 40 % identity NCBI Description (AC003680) putative DNA binding protein [Arabidopsis thaliana] 5808 Seq. No. 5448 6.R1040 Contig ID LIB3170-017-Q1-K1-H11 5'-most EST Method BLASTX NCBI GI a2580440 BLAST score 300 E value 4.0e-27 Match length 74 % identity 74 NCBI Description (D87261) PCF2 [Oryza sativa] Seq. No. 5809 5449 1.R1040 Contig ID 5'-most EST LIB3028-001-Q1-B1-E11 Method BLASTX NCBI GI g2245108 BLAST score 382 E value 1.0e-36 Match length 157 % identity 52 NCBI Description (Z97343) EREBP-4 homolog [Arabidopsis thaliana] Seq. No. 5810 Contig ID 5451 1.R1040 5'-most EST ssr700559029.h1 Seq. No. 5811 5451 2.R1040 Contig ID 5'-most EST LIB3028-001-Q1-B1-B6 5812 Seq. No. 5452 1.R1040 Contig ID 5'-most EST LIB3030-005-Q1-B1-C12 Method BLASTX NCBI GI g2661840 BLAST score 1216 E value 1.0e-134 Match length 340 % identity NCBI Description (Y15430) adenosine kinase [Physcomitrella patens] 5813 Seq. No. 5452 2.R1040 Contig ID

Contig ID 5452_2.R1040 5'-most EST xpa700797523.h1

Method BLASTX
NCBI GI g2244759
BLAST score 2192



E value 0.0e+00 Match length 489 % identity 79

NCBI Description (Z97335) selenium-binding protein [Arabidopsis thaliana]

Seq. No. 5814

Contig ID 5452_6.R1040 5'-most EST asn701131792.h1

Method BLASTX
NCBI GI g2244760
BLAST score 616
E value 4.0e-64
Match length 151
% identity 73

NCBI Description (Z97335) selenium-binding protein [Arabidopsis thaliana]

Seq. No. 5815

Contig ID 5452_7.R1040 5'-most EST hyd700729029.h1

Method BLASTX
NCBI GI g2661840
BLAST score 329
E value 1.0e-30
Match length 98
% identity 59

NCBI Description (Y15430) adenosine kinase [Physcomitrella patens]

Seq. No. 5816

Contig ID 5454_1.R1040 5'-most EST dpv701099580.h1

Seq. No. 5817

Contig ID 5454_2.R1040

5'-most EST LIB3028-001-Q1-B1-C1

Seq. No. 5818

Contig ID 5454_3.R1040 5'-most EST zzp700835538.h1

Seq. No. 5819

Contig ID 5455 1.R1040

5'-most EST LIB3049-055-Q1-E1-B8

Method BLASTX
NCBI GI g3876299
BLAST score 325
E value 7.0e-30
Match length 135
% identity 48

NCBI Description (Z71180) similar to BPTI/KUNITZ inhibitor domain; cDNA EST EMBL:D68293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk249e6.5 comes from this gene; cDNA EST yk448h4.3 comes from this gene [Caenorhabdi... >gi_3880760_emb_CAA16311_ (AL021474) similar to BPTI/KUNITZ inhibitor domain; cDNA EST EMBL:D68293 comes from this

inhibitor domain; cDNA EST EMBL: D68293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk249e6.5 comes from this gene; cDNA EST yk448h4.3 comes

from this gene [Caenorhab



```
5820
 Seq. No.
                    5456 1.R1040
 Contig ID
                    leu701156483.hl
 5'-most EST
                    BLASTX
 Method
                    g3928090
 NCBI GI
                    318
 BLAST score
                    2.0e-29
 E value
 Match length
                    104
                     60
 % identity
                    (AC005770) putative MTN3 protein [Arabidopsis thaliana]
 NCBI Description
                     5821
 Seq. No.
 Contig ID
                     5458 1.R1040
                     zhf700958912.h1
  5'-most EST
                     BLASTX
 Method
 NCBI GI
                     g3420055
 BLAST score
                     864
                     8.0e-93
 E value
 Match length
                     184
                     88
  % identity
                     (AC004680) cyclophilin [Arabidopsis thaliana]
 NCBI Description
  Seq. No.
                     5460 1.R1040
  Contig ID
                     kmv700738102.hl
  5'-most EST
                     5823
  Seq. No.
                     5460 2.R1040
  Contig ID
                     LIB3051-083-Q1-K1-F6
  5'-most EST
                     BLASTX
  Method
                     g1488647
  NCBI GI
                     2033
  BLAST score
                     0.0e + 00
  E value
  Match length
                     609
  % identity
                     67
                     (X99937) RNA helicase [Spinacia oleracea]
  NCBI Description
                     5824
  Seq. No.
                     5460 3.R1040
  Contig ID
                     uC-gmropic096e01b1
  5'-most EST
                     5825
  Seq. No.
                     5467 1.R1040
Contig ID
                     LIB3\overline{0}49-003-Q1-E1-B10
  5'-most EST
                     BLASTN
  Method
                     q458337
  NCBI GI
  BLAST score
                     539
                     0.0e + 00
  E value
                     1127
  Match length
                     89
  % identity
                     Vigna radiata clone pVR-ACO1
  NCBI Description
                     1-aminocyclopropane-1-carboxylate oxidase homolog mRNA,
                     complete cds
```

1105

5826

5467 2.R1040

Seq. No.

Contig ID



leu701146339.hl 5'-most EST BLASTN Method g458337 NCBI GI BLAST score 256 1.0e-142 E value 442 Match length 90 % identity Vigna radiata clone pVR-ACO1 NCBI Description 1-aminocyclopropane-1-carboxylate oxidase homolog mRNA, complete cds 5827 Seq. No. 5468 1.R1040 Contig ID ssr700555552.hl 5'-most EST Method BLASTX NCBI GI q3785983 BLAST score 401 6.0e - 39E value 117 Match length 62 % identity (AC005560) hypothetical protein [Arabidopsis thaliana] NCBI Description 5828 Seq. No. 5469_1.R1040 Contig ID 5'-most EST jC-gmst02400032b01a1 BLASTX Method q2623296 NCBI GI 677 BLAST score E value 4.0e-71 232 Match length % identity 65 (AC002409) hypothetical protein [Arabidopsis thaliana] NCBI Description 5829 Seq. No. Contig ID 5469 3.R1040 5'-most EST LIB3106-100-Q1-K1-B8 Seq. No. 5830 5471 1.R1040 Contig ID LIB3051-063-Q1-K1-D9 5'-most EST Method BLASTX g2462762 NCBI GI 959 BLAST score 1.0e-104 E value Match length 288 64 % identity (AC002292) Highly similar to auxin-induced protein NCBI Description (aldo/keto reductase family) [Arabidopsis thaliana] 5831 Seq. No. 5472 1.R1040 Contig ID LIB3028-001-Q1-B1-A11 5'-most EST

Method BLASTX
NCBI GI g4262149
BLAST score 1131
E value 1.0e-124
Match length 268



% identity 73
NCBI Description (AC005275) putative xyloglucan endotransglycosylase
[Arabidopsis thaliana]
Seq. No. 5832
Contig ID 5472 2.R1040
5'-most EST g4260368

Contig ID 5472_2.R1040
5'-most EST g4260368
Method BLASTX
NCBI GI g4262149
BLAST score 311
E value 2.0e-28
Match length 93
% identity 57

NCBI Description (AC005275) putative xyloglucan endotransglycosylase

[Arabidopsis thaliana]

Seq. No. 5833

Contig ID 5472_3.R1040 5'-most EST zhf700953943.h1

Method BLASTX
NCBI GI g4262149
BLAST score 543
E value 9.0e-56
Match length 170
% identity 59

NCBI Description (AC005275) putative xyloglucan endotransglycosylase

[Arabidopsis thaliana]

Seq. No. 5834

Contig ID 5473 1.R1040

5'-most EST LIB3028-001-Q1-B1-B10

Method BLASTX
NCBI GI g2654559
BLAST score 314
E value 1.0e-28
Match length 152
% identity 39

NCBI Description (AF006621) embryonic lung protein [Homo sapiens]

Seq. No. 5835

Contig ID 5475_1.R1040 5'-most EST zsg701119377.h1

Seq. No. 5836

Contig ID 5480 1.R1040

5'-most EST LIB3028-001-Q1-B1-E10

Method BLASTX
NCBI GI g4539422
BLAST score 581
E value 6.0e-60
Match length 134
% identity 76

NCBI Description (ALO49171) putative protein [Arabidopsis thaliana]

Seq. No.

Contig ID 5480_2.R1040

5'-most EST LIB3049-013-Q1-E1-G1



```
BLASTX
Method
NCBI GI
                  g4539422
                  535
BLAST score
                  5.0e-55
E value
                  110
Match length
                  83
% identity
                  (AL049171) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   5838
                   5481 1.R1040
Contig ID
                  LIB3028-001-Q1-B1-E12
5'-most EST
                  BLASTX
Method
                   g1213629
NCBI GI
                   771
BLAST score
                   4.0e-82
E value
                   238
Match length
                   60
% identity
                   (X95991) pectinesterase [Prunus persica]
NCBI Description
                   5839
Seq. No.
                   5482 1.R1040
Contig ID
                   jC-qmst02400072a09d1
5'-most EST
                   BLASTX
Method
                   q2108252
NCBI GI
                   317
BLAST score
                   3.0e-29
E value
                   102
Match length
                   34
% identity
                   (Y10228) P-glycoprotein-2 [Arabidopsis thaliana]
NCBI Description
                   >qi 2108254 emb CAA71276_ (Y10227) P-glycoprotein-2
                   [Arabidopsis thaliana] >gi_4538925_emb_CAB39661.1
                   (AL049483) P-glycoprotein-2 (pgp2) [Arabidopsis thaliana]
                   5840
Seq. No.
                   5486 1.R1040
Contig ID
                   LIB3028-001-Q1-B1-G10
5'-most EST
Method
                   BLASTX
NCBI GI
                   q132724
BLAST score
                   305
                   2.0e-27
E value
                   118
Match length
                   54
% identity
                   50S RIBOSOMAL PROTEIN L18 >gi_71259_pir__R5BS8F ribosomal
NCBI Description
                   protein L18 - Bacillus stearothermophilus
                   5841
Seq. No.
                   5487 1.R1040
Contig ID
5'-most EST
                   gsv701056432.h1
Method
                   BLASTN
                   g20866
NCBI GI
BLAST score
                   196
                   1.0e-105
E value
```

Match length 585 % identity 85

P.sativum mRNA for plastid ribosomal protein CL15 NCBI Description

5842 Seq. No.



Contig ID 5487_2.R1040 5'-most EST LIB3028-001-Q1-B1-G11

Seq. No. 5843

Contig ID 5487 3.R1040 5'-most EST pcp700991074.h1

Method BLASTN
NCBI GI g20866
BLAST score 53
E value 4.0e-21
Match length 161
% identity 83

NCBI Description P.sativum mRNA for plastid ribosomal protein CL15

Seq. No.

Contig ID 5497 1.R1040

5'-most EST jC-gmro02910072a11d1

5844

Method BLASTX
NCBI GI g4204294
BLAST score 240
E value 4.0e-20
Match length 82
% identity 62

NCBI Description (AC003027) lcl prt seq No definition line found

[Arabidopsis thaliana]

Seq. No. 5845

Contig ID 5498 1.R1040

5'-most EST LIB3049-011-Q1-E1-D3

Seq. No. 5846

Contig ID 5499 1.R1040

5'-most EST LIB3109-056-Q1-K1-E12

Method BLASTX
NCBI GI g1709358
BLAST score 1489
E value 1.0e-166
Match length 441

% identity 64

NCBI Description NUCLEOSIDE-TRIPHOSPHATASE (NUCLEOSIDE TRIPHOSPHATE

PHOSPHOHYDROLASE) (NTPASE) >gi_629638_pir__\$48859

nucleoside triphosphatase - garden pea

>gi_2129890_pir__S65147 nucleoside triphosphatase

precursor, chromatin-associated - garden pea

>gi 563612 emb CAA83655 (Z32743) nucleoside triphosphatase

[Pisum sativum] >gi_4519173_dbj_BAA75506.1_ (AB022319) nucleoside triphosphatase (NTPase) [Pisum sativum]

Seq. No. 5847

Contig ID 5501_1.R1040 5'-most EST ncj700975782.h1

Method BLASTX
NCBI GI g2982303
BLAST score 387
E value 2.0e-37
Match length 83
% identity 83



NCBI Description (AF051236) hypothetical protein [Picea mariana]

Seq. No.

5848

Contig ID

5508 1.R1040

5'-most EST

LIB3028-005-Q1-B1-C12

Method NCBI GI BLASTX

BLAST score

g2398829

E value

640 1.0e-66

Match length

239 29

% identity

NCBI Description (Y11220) mitochondrial uncoupling protein [Solanum

tuberosum]

Seq. No.

5849

Contig ID

5508 2.R1040

5'-most EST Method

LIB3040-031-Q1-E2-H5

NCBI GI

BLASTX

BLAST score

g2398829 418

E value

7.0e-41

Match length

111

% identity

NCBI Description (Y11220) mitochondrial uncoupling protein [Solanum

tuberosum]

Seq. No.

5850

Contig ID

5510 1.R1040

5'-most EST

LIB3028-005-Q1-B1-D4

Method

BLASTX g2583112

NCBI GI BLAST score

315

E value

9.0e-29

Match length

142

% identity

NCBI Description (AC002387) putative PD1-like DNA-binding protein

[Arabidopsis thaliana]

Seq. No.

5851

Contig ID 5'-most EST 5510 2.R1040 $hyd7\overline{0}0726712.h1$

Method

BLASTX

NCBI GI

g3668079

BLAST score

177

E value

8.0e-13

Match length

% identity

62

NCBI Description (AC004667) hypothetical protein [Arabidopsis thaliana]

Seq. No.

5852

Contig ID

5512 1.R1040

5'-most EST

LIB3092-055-Q1-K1-G10

Seq. No.

5853

Contig ID

5517 1.R1040

5'-most EST

LIB3028-006-Q1-B1-D5



```
Seq. No.
                   5854
Contig ID
                  5520 1.R1040
5'-most EST
                  q5605979
Method
                  BLASTX
NCBI GI
                  g4006829
                  675
BLAST score
                  1.0e-70
E value
Match length
                  304
% identity
                  56
                 (AC005970) putative protein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  5855
                  5520 2.R1040
Contig ID
5'-most EST
                  jC-qmro02800037a12a1
Method
                  BLASTX
NCBI GI
                  g3461835
BLAST score
                  856
E value
                  6.0e-92
Match length
                  218
                  74
% identity
                  (AC005315) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                  >qi 3927840 (AC005727) putative protein kinase [Arabidopsis
                  thaliana]
Seq. No.
                  5856
Contig ID
                   5520 3.R1040
5'-most EST
                  smc700746635.hl
Seq. No.
                  5857
                  5521 1.R1040
Contig ID
5'-most EST
                  trc700563573.h1
Method
                  BLASTX
NCBI GI
                  g2191150
BLAST score
                  173
                   4.0e-12
E value
Match length
                   98
% identity
                  14
NCBI Description
                   (AF007269) similar to mitochondrial carrier family
                   [Arabidopsis thaliana]
                   5858
Seq. No.
```

5523 1.R1040 Contig ID

5'-most EST LIB3028-007-Q1-B1-F11

Seq. No. 5859

5528 1.R1040 Contig ID

5'-most EST LIB3106-101-Q1-K1-E2

Seq. No.

Contig ID 5528 2.R1040 5'-most EST leu701152989.h1

5861 Seq. No.

Contig ID 5529 1.R1040

5'-most EST LIB3072-012-Q1-E1-G3

Method BLASTX NCBI GI g548774



BLAST score 1028 E value 1.0e-112 Match length 240 % identity 82

NCBI Description 60S RIBOSOMAL PROTEIN L7A >gi_542158_pir__S38360 ribosomal

protein L7a - rice >gi_303855_dbj_BAA02156_ (D12631)

ribosomal protein L7A [Oryza sativa]

Seq. No. 5862

Contig ID 5536_1.R1040 5'-most EST asn701135418.h1

Method BLASTX
NCBI GI g4234941
BLAST score 848
E value 7.0e-99
Match length 210
% identity 89

NCBI Description (AF097938) cytosolic phosphoglucomutase [Populus tremula x

Populus tremuloides]

Seq. No. 5863

Contig ID 5540 1.R1040

5'-most EST jC-gmf102220094d09a1

Method BLASTX
NCBI GI g2102691
BLAST score 412
E value 4.0e-40
Match length 125
% identity 65

NCBI Description (U64817) fructokinase [Lycopersicon esculentum]

Seq. No. 5864

Contig ID 5543_1.R1040

5'-most EST jC-gmst02400041g07a1

Method BLASTX
NCBI GI g123601
BLAST score 766
E value 2.0e-81
Match length 242
% identity 62

NCBI Description HEAT SHOCK 70 KD PROTEIN >gi_99913_pir__S14992 heat shock

protein, 70K - soybean >gi_18663_emb_CAA44620_ (X62799)

Heat Shock 70kD protein [Glycine max]

Seq. No. 5865

Contig ID 5547 1.R1040

5'-most EST LIB3028-009-Q1-B1-B4

Method BLASTX
NCBI GI g2827704
BLAST score 218
E value 2.0e-17
Match length 82
% identity 54

NCBI Description (AL021684) LRR-like protein [Arabidopsis thaliana]

Seq. No. 5866

Contig ID 5548 1.R1040



```
5'-most EST
                  uC-gmropic012c05b1
Method
                  BLASTX
                  g2464905
NCBI GI
                  295
BLAST score
                  2.0e-26
E value
Match length
                  148
                  47
% identity
NCBI Description (299708) minor allergen [Arabidopsis thaliana]
                  5867
Seq. No.
                  5548 2.R1040
Contig ID
5'-most EST
                  LIB3028-009-Q1-B1-C1
Method
                  BLASTN
                  q4006885
NCBI GI
BLAST score
                  34
                  1.0e-09
E value
Match length
                  78
                  92
% identity
NCBI Description
                  Arabidopsis thaliana DNA chromosome 4, ESSA I AP2 contig
                  fragment No
Seq. No.
                  5868
                  5550 1.R1040
Contig ID
5'-most EST
                  uxk700667830.hl
Method
                  BLASTX
NCBI GI
                  g3834322
                  274
BLAST score
E value
                  8.0e-24
Match length
                  189
% identity
                  39
NCBI Description (AC005679) EST gb R30300 comes from this gene. [Arabidopsis
                  thaliana]
Seq. No.
                  5869
Contig ID
                  5550 2.R1040
5'-most EST
                  LIB3049-009-Q1-E1-A1
Method
                  BLASTN
                  g4006885
NCBI GI
BLAST score
                  62
                  5.0e-26
E value
Match length
                  150
% identity
                  85
NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I AP2 contig
                  fragment No
Seq. No.
                   5870
Contig ID
                   5552 1.R1040
5'-most EST
                  LIB3051-006-Q1-E1-F9
Method
                  BLASTX
                  g4006867
NCBI GI
BLAST score
                  214
                  5.0e-17
E value
                  53
Match length
                  77
% identity
```

Seq. No. 5871

NCBI Description

(Z99707) putative protein [Arabidopsis thaliana]



```
5552 2.R1040
Contig ID
                   LIB3051-091-Q1-K1-E9
5'-most EST
Seq. No.
                   5554 1.R1040
Contig ID
                   LIB3028-009-Q1-B1-E4
5'-most EST
                   BLASTX
Method
                   q4469003
NCBI GI
BLAST score
                   367
                   2.0e-34
E value
Match length
                   102
                   63
% identity
                  (AL035602) putative protein [Arabidopsis thaliana]
NCBI Description
                   5873
Seq. No.
                   5554 2.R1040
Contig ID
                   wrq700788935.h2
5'-most EST
                   5874
Seq. No.
                   5556 1.R1040
Contig ID
                   q429\overline{1}358
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2462837
BLAST score
                   262
                   2.0e-22
E value
Match length
                   184
% identity
                   36
                   (AF000657) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   5875
Seq. No.
                   5557 1.R1040
Contig ID
                   LIB3138-044-Q1-N1-A10
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3913711
BLAST score
                   2382
E value
                   0.0e + 00
                   495
Match length
                   93
% identity
                   NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE
NCBI Description
                    (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE
                   DEHYDROGENASE) (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE
                    (NADP+)) (TRIOSEPHOSPHATE DEHYDROGENASE)
                   >gi 4379384 emb_CAA53076_ (X75327)
                   glyceraldehyde-3-phosphate dehydrogenase
                    (nonphosphorylating, NADP+) [Pisum sativum]
                   5876
Seq. No.
                   5557 5.R1040
Contig ID
                   hyd7\overline{0}0726390.h1
5'-most EST
                   BLASTN
Method
                   g496493
NCBI GI
```

73 BLAST score

5.0e-33 E value Match length 226 % identity 88

P.sativum (Rosakrone) mRNA for nonphosphorylating, NCBI Description NADP-specific, glyceraldehyde-3-phosphate dehydrogenase



```
Seq. No.
                  5877
                  5562 1.R1040
Contig ID
5'-most EST
                  LIB3039-040-Q1-E1-C8
Method
                  BLASTN
                  g3776081
NCBI GI
BLAST score
                  217
E value
                  1.0e-118
Match length
                  353
                  90
% identity
NCBI Description Medicago truncatula mRNA for MtN30 gene, partial
Seq. No.
                  5878
                  5562 4.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy104b02b1
Method
                  BLASTX
NCBI GI
                  g3892712
BLAST score
                  338
E value
                  1.0e-31
Match length
                  76
% identity
                  84
NCBI Description
                  (AL033545) adenine phosphoribosyltransferase (EC 2.4.2.7) -
                  like protein [Arabidopsis thaliana]
                  5879
Seq. No.
                  5562 5.R1040
Contig ID
5'-most EST
                  LIB3028-009-Q1-B1-H2
Method
                  BLASTN
                  q3776081
NCBI GI
BLAST score
                  147
                  5.0e-77
E value
Match length
                  231
% identity
NCBI Description Medicago truncatula mRNA for MtN30 gene, partial
                  5880
Seq. No.
Contig ID
                  5564_1.R1040
5'-most EST
                  jC-gmf102220051h02a1
Method
                  BLASTX
NCBI GI
                  g3759184
BLAST score
                  651
                  4.0e-68
E value
Match length
                  198
% identity
                  63
NCBI Description (AB018441) phi-1 [Nicotiana tabacum]
                  5881
Seq. No.
Contig ID
                  5568 1.R1040
5'-most EST
                  LIB3028-010-Q1-B1-B2
```

Seq. No.

5882

Contig ID

5570 1.R1040

5'-most EST

LIB3028-010-Q1-B1-B4

Method BLASTX
NCBI GI g1168750
BLAST score 269
E value 2.0e-23



Match length 157 % identity

CALCINEURIN B SUBUNIT (PROTEIN PHOSPHATASE 2B REGULATORY NCBI Description SUBUNIT) >gi 458230 (U04380) calcineurin B [Naegleria

gruberi]

Seq. No.

5883

Contig ID 5'-most EST

5573 1.R1040 ncj700983528.h1

Method BLASTX NCBI GI g557882 BLAST score 361 E value 5.0e - 34Match length 183 % identity 45

NCBI Description

(U13923) putative serine/threonine protein kinase; similar to product encoded by Lycopersicon pimpinellifolium Pto gene, GenBank Accession Number U02271; Fen is a member of the Pto gene family [Solanum pimpinellifolium] >gi_1809263 (U59317) serine/threonine protein kinase Fen [Lycopersicon pimpinellifolium] >gi 1096880 prf 2112354B Fen gene [Lycopersicon esculentum] >gi 1098334 prf 2115395A Fen

gene [Lycopersicon esculentum]

Seq. No.

5884

Contig ID 5'-most EST 5573 2.R1040 zsg701127563.h1

Seq. No.

5885

Contig ID

5575 1.R1040

5'-most EST uC-gmropic090e01b1

Seq. No.

5886

Contig ID 5'-most EST

5576 1.R1040 zzp700831186.h1

Method BLASTX NCBI GI g2501102 BLAST score 212 E value 6.0e-17 Match length 100 % identity 49

NCBI Description

SYNTAXIN-RELATED PROTEIN KNOLLE >gi 1184165 (U39451)

syntaxin-related [Arabidopsis thaliana] >gi_1184167 (U39452) syntaxin-related [Arabidopsis thaliana]

>gi 3063443 (AC003981) F22013.4 [Arabidopsis thaliana] >gi 1587182 prf 2206310A syntaxin-related protein

[Arabidopsis thaliana]

Seq. No.

5887

Contig ID

5577 1.R1040

5'-most EST

LIB3028-010-Q1-B1-E1

Method BLASTX NCBI GI q4204315 BLAST score 404 E value 9.0e-42Match length 148 % identity 59

Seq. No.

Contig ID



```
(AC003027) Unknown protein [Arabidopsis thaliana]
NCBI Description
                   5888
Seq. No.
                   5577 2.R1040
Contig ID
                  LIB3138-015-Q1-N2-F2
5'-most EST
                   BLASTX
Method
                   q4204315
NCBI GI
BLAST score
                   402
                   3.0e-39
E value
                   130
Match length
                                 54 1
                   56
% identity
                   (AC003027) Unknown protein [Arabidopsis thaliana]
NCBI Description
                   5889
Seq. No.
                   5577 3.R1040
Contig ID
5'-most EST
                   bth700848565.h1
                   BLASTX
Method
                   q4204315
NCBI GI
                   425
BLAST score
                   7.0e-42
E value
Match length
                   138
                   55
% identity
                  (AC003027) Unknown protein [Arabidopsis thaliana]
NCBI Description
                   5890
Seq. No.
                   5581 1.R1040
Contig ID
                   LIB3106-063-Q1-K1-B8
5'-most EST
                   BLASTX
Method
                   q3323581
NCBI GI
                   409
BLAST score
                   6.0e-40
E value
                   101
Match length
                   83
% identity
                   (AF057281) two-component response regulator homolog
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   5891
                   5581 3.R1040
Contig ID
                   LIB3028-010-Q1-B1-E5
5'-most EST
                   5892
Seq. No.
                   5586 1.R1040
Contig ID
                   LIB3028-010-Q1-B1-G1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4006827
BLAST score
                   433
E value
                   4.0e-43
                   193
Match length
                   47
% identity
                   (AC005970) subtilisin-like protease [Arabidopsis thaliana]
NCBI Description
                   5893
Seq. No.
                   5588 1.R1040
Contig ID
                   LIB3028-010-Q1-B1-G4
5'-most EST
                   5894
```

5589_1.R1040



5'-most EST LIB3028-010-Q1-B1-H1

Method BLASTX
NCBI GI 9710626
BLAST score 303
E value 4.0e-27
Match length 87
% identity 72

NCBI Description (D30719) ERD15 protein [Arabidopsis thaliana] >gi_3241941

(AC004625) dehydration-induced protein ERD15 [Arabidopsis thaliana] >gi_3894181 (AC005662) ERD15 protein [Arabidopsis

thaliana]

Seq. No. 5895

Contig ID 5591 1.R1040

5'-most EST LIB3094-054-Q1-K1-A3

Method BLASTX
NCBI GI g2764941
BLAST score 323
E value 1.0e-29
Match length 89
% identity 61

NCBI Description (X98255) transcriptionally stimulated by gibberellins;

expressed in meristematic region, and style [Arabidopsis

thaliana]

Seq. No. 5896

Contig ID 5591_2.R1040

5'-most EST LIB3094-078-Q1-K1-E1

Method BLASTX
NCBI GI g2764941
BLAST score 374
E value 8.0e-36
Match length 89
% identity 69

NCBI Description (X98255) transcriptionally stimulated by gibberellins;

expressed in meristematic region, and style [Arabidopsis

thaliana]

Seq. No. 5897

Contig ID 5592_1.R1040

5'-most EST LIB3028-010-Q1-B1-H4

Seq. No. 5898

Contig ID 5593 1.R1040

5'-most EST LIB3028-011-Q1-B1-A1

Method BLASTX
NCBI GI g2244971
BLAST score 166
E value 3.0e-11
Match length 39
% identity 77

NCBI Description (Z97340) hypothetical protein [Arabidopsis thaliana]

Seq. No. 5899

Contig ID 5594 1.R1040

5'-most EST LIB3049-011-Q1-E1-B2



Seq. No. 5900

5595 1.R1040 Contig ID

5'-most EST LIB3052-007-Q1-B1-A2

Seq. No.

5901

Contig ID

5597 1.R1040

5'-most EST

LIB3072-012-Q1-E1-G8

Method NCBI GI BLASTX g2047324

BLAST score E value

696 3.0e-73

Match length

157

% identity

78

NCBI Description (U80192) HAL3 homolog [Arabidopsis thaliana]

Seq. No.

5902

Contig ID

5597 2.R1040

5'-most EST

LIB3051-028-Q1-K1-D1

Method NCBI GI BLASTX

BLAST score

g4106515 648

E value

1.0e-80

Match length

183

% identity

80

NCBI Description (AF092743) CAK associated cyclinH homolog [Populus tremula

....

x Populus tremuloides]

Seq. No.

5903

Contig ID

5597 3.R1040

5'-most EST

LIB3093-058-Q1-K1-A11

Seq. No.

5904

Contig ID

5597 4.R1040

5'-most EST

LIB3109-047-Q1-K1-A10

Seq. No.

5905

Contig ID

5597 5.R1040

5'-most EST

g5126313

Seq. No.

5906

Contig ID

5597 6.R1040

5'-most EST

leu701147304.h1

Seq. No.

5907

Contig ID

5597 7.R1040

5'-most EST

LIB3093-012-Q1-K1-D9

Method

BLASTX

NCBI GI

q549706 193

BLAST score

E value Match length 2.0e-14

% identity

113

NCBI Description KTI12 PROTEIN >gi 539197 pir S37937 KTI12 protein - yeast

(Saccharomyces cerevisiae) >gi 486185 emb CAA81950

(Z28110) ORF YKL110c [Saccharomyces cerevisiae]

>gi 536816 emb CAA54646 (X77511) KTI12 [Saccharomyces

cerevisiae]



```
5908
Seq. No.
                   5597 8.R1040
Contig ID
5'-most EST
                   yza700764032.h1
                   5909
Seq. No.
                   5597 9.R1040
Contig ID
5'-most EST
                   zsg701129470.h1
Method
                   BLASTN
NCBI GI
                   q4519195
BLAST score
                   36
                   2.0e-10
E value
Match length
                   172
% identity
                   85
                  Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:
NCBI Description
                   MQC12, complete sequence
Seq. No.
                   5910
                   5597 11.R1040
Contig ID
5'-most EST
                   LIB3050-020-Q1-K1-C4
Seq. No.
                   5911
Contig ID
                   5597 12.R1040
5'-most EST
                   uC-gmrominsoy201a07b1
                   5912
Seq. No.
Contig ID
                   5597 14.R1040
5'-most EST
                   uaw700661065.hl
Seq. No.
                   5913
Contig ID
                   5597 20.R1040
5'-most EST
                   dpv701098531.hl
                   5914
Seq. No.
                   5597 21.R1040
Contig ID
5'-most EST
                   ncj7\overline{0}0979023.h1
                   5915
Seq. No.
                   5603 1.R1040
Contig ID
5'-most EST
                   gsv701055160.hl
Seq. No.
                   5916
                   5605 1.R1040
Contig ID
5'-most EST
                   LIB3028-012-Q1-B1-A1
Seq. No.
                   5917
                   5608 1.R1040
Contig ID
5'-most EST
                   ncj700981742.hl
                   BLASTX
Method
```

Method BLASTX
NCBI GI g2497702
BLAST score 137
E value 8.0e-15
Match length 118
% identity 47

NCBI Description OUTER MEMBRANE LIPOPROTEIN BLC PRECURSOR

>gi_2121019_pir__I40710 outer membrane lipoprotein Citrobacter freundii >gi_717136 (U21727) lipocalin



```
precursor [Citrobacter freundii]
Seq. No.
                  5918
Contig ID
                  5611 1.R1040
                  ncj700988402.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3115204
BLAST score
                  334
E value
                  2.0e-30
Match length
                  304
% identity
                  22
                  (Y12059) strong homology to human RING3 sequence [Homo
NCBI Description
                  sapiens]
Seq. No.
                  5919
                  5611 2.R1040
Contig ID
5'-most EST
                  g5677950
                  5920
Seq. No.
                  5611 3.R1040
Contig ID
5'-most EST
                  LIB3092-030-Q1-K1-G11
Seq. No.
                  5921
                  5612 1.R1040
Contig ID
                  zhf700958085.h1
5'-most EST
Seq. No.
                  5922
Contig ID
                  5613 1.R1040
5'-most EST
                  LIB3028-013-Q1-B1-B7
Method
                  BLASTX
NCBI GI
                  g4467146
BLAST score
                  679
                  3.0e-71
E value
Match length
                  254
                  53
% identity
                  (AL035540) galactosidase like protein [Arabidopsis
NCBI Description
                  thaliana]
                   5923
Seq. No.
Contig ID
                   5615 1.R1040
5'-most EST
                  kl1701215452.h1
Method
                  BLASTX
NCBI GI
                  g3334320
BLAST score
                  1131
                  1.0e-124
E value
                   273
Match length
% identity
                  80
                  40S RIBOSOMAL PROTEIN SA (P40) >gi 2444420 (AF020553)
NCBI Description
                   ribosome-associated protein p40 [GTycine max]
                   5924
Seq. No.
                   5617 1.R1040
```

Contig ID

5'-most EST jC-gmro02910065c04d1

BLASTX Method NCBI GI q4572673 BLAST score 551 E value 2.0e-56



Match length % identity

(AC006954) putative sarcosine oxidase [Arabidopsis NCBI Description

thaliana]

5925 Seq. No.

5618 1.R1040 Contig ID 5'-most EST rry700808482.h1

Method BLASTX NCBI GI g4263787 405 BLAST score 2.0e-39 E value 120 Match length 72 % identity

(AC006068) unknown protein [Arabidopsis thaliana] NCBI Description

5926 Seq. No.

5623 1.R1040 Contig ID $uaw7\overline{0}0661344.h1$ 5'-most EST

Method BLASTX NCBI GI g4098678 BLAST score 300 2.0e-26 E value 305 Match length 19 % identity

(U80821) C1 transcription factor [Mus musculus] NCBI Description

5927 Seq. No.

5624 1.R1040 Contig ID art700605411.h2 5'-most EST

BLASTX Method q730165 NCBI GI 248 BLAST score 8.0e-21 E value 80 Match length 60 % identity

EARLY NODULIN 93 (N-93) >gi_486679_pir__S34801 nodulin NCBI Description (clone GmN93) - soybean >gi_218262_dbj_BAA02724_ (D13506)

early nodulin [Glycine max] >gi 3763851 dbj_BAA33816

(AB018378) early nodulin [Glycine max] >gi_447138_prf__1913422D nodulin [Glycine max]

5928 Seq. No.

5628 1.R1040 Contig ID

LIB3028-014-Q1-B1-C4 5'-most EST

Method BLASTX a418754 NCBI GI 224 BLAST score 4.0e-18 E value 110 Match length % identity

catechol oxidase (EC 1.10.3.1) precursor - fava bean NCBI Description

Seq. No. 5929

5629 1.R1040 Contig ID

LIB3028-014-Q1-B1-D12 5'-most EST



```
5930
Seq. No.
                   5629 2.R1040
Contig ID
                   sat701014978.hl
5'-most EST
                   5931
Seq. No.
                   5633 1.R1040
Contig ID
5'-most EST
                   LIB3028-014-Q1-B1-F12
Method
                   BLASTX
                   g4336205
NCBI GI
                   359
BLAST score
                   9.0e - 34
E value
                   191
Match length
                   42
% identity
                   (AF077372) cytochrome b5 reductase [Zea mays]
NCBI Description
Seq. No.
                   5633 2.R1040
Contig ID
5'-most EST
                   LIB3040-020-Q1-E1-H2
                   5933
Seq. No.
                   5641 1.R1040
Contig ID
                   k117\overline{0}1209531.h1
5'-most EST
                   5934
Seq. No.
                   5641 2.R1040
Contig ID
                   LIB3028-016-Q1-B1-D4
5'-most EST
                   5935
Seq. No.
                   5641 3.R1040
Contig ID
                   sat701002867.hl
5'-most EST
                   5936
Seq. No.
                   5641 4.R1040
Contig ID
                   epx701109189.h1
5'-most EST
                   5937
Seq. No.
                   5642 1.R1040
Contig ID
                   pmv700892068.h1
5'-most EST
Method
                   BLASTX
                   g4056432
NCBI GI
BLAST score
                   714
                   3.0e-75
E value
Match length
                   221
% identity
                   (AC005990) Similar to gi 2245014 glucosyltransferase
NCBI Description
                   homolog from Arabidopsis thaliana chromosome 4 contig
                   gb_Z97341. ESTs gb_T20778 and gb_AA586281 come from this
                   gene. [Arabidopsis thaliana]
                   5938
Seq. No.
```

Contig ID 5642_2.R1040 5'-most EST hyd700726120.h1

Method BLASTX
NCBI GI g4263795
BLAST score 389
E value 1.0e-37
Match length 126



```
% identity
                  (AC006068) putative glucosyltransferase [Arabidopsis
NCBI Description
                  thaliana]
                  5939
Seq. No.
                  5644 1.R1040
Contig ID
                  LIB3028-016-Q1-B1-H3
5'-most EST
Method
                  BLASTX
                  q3123295
NCBI GI
                  383
BLAST score
                   1.0e-36
E value
                  136
Match length
                   39
% identity
                  CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi 2583169
NCBI Description
                   (AF026473) calmodulin-related protein [Arabidopsis
                   thaliana]
                   5940
Seq. No.
                   5648 1.R1040
Contig ID
                   LIB3028-047-Q1-B1-G10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4567260
                   740
BLAST score
                   1.0e-78
E value
                   166
Match length
                   84
% identity
                   (AC006841) putative NADPH dependent mannose 6-phosphate
NCBI Description
                   reductase [Arabidopsis thaliana]
                   5941
Seq. No.
                   5648 2.R1040
Contig ID
                   LIB3167-049-P1-K1-H1
5'-most EST
                   BLASTX
Method
                   a4567260
NCBI GI
                   383
BLAST score
                   2.0e-65
E value
                   163
Match length
                   82
% identity
                   (AC006841) putative NADPH dependent mannose 6-phosphate
NCBI Description
                   reductase [Arabidopsis thaliana]
                   5942
Seq. No.
Contig ID
                   5649 1.R1040
 5'-most EST
                   pmv700894289.hl
                   BLASTN
Method
                   q3021347
NCBI GI
 BLAST score
                   160
```

2.0e-84 E value Match length 555 83 % identity

NCBI Description Cicer arietinum mRNA for cytoplasmic ribosomal protein L18

Seq. No. 5943 5651 1.R1040 Contig ID g4395973 5'-most EST BLASTX Method q3885338

NCBI GI

```
BLAST score
                   1.0e-12
E value
                   168
Match length
                   27
% identity
                   (AC005623) unknown protein [Arabidopsis thaliana]
NCBI Description
                   5944
Seq. No.
                   5651 2.R1040
Contig ID
5'-most EST
                   LIB3051-025-Q1-K1-D8
                   5945
Seq. No.
                   5652 1.R1040
Contig ID
5'-most EST
                   LIB3028-018-Q1-B1-A11
                   5946
Seq. No.
                   5653 1.R1040
Contig ID
                   jC-gmro02800031d05d1
5'-most EST
                   5947
Seq. No.
                   5657 1.R1040
Contig ID
                   LIB3107-079-Q1-K1-F10
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2388585
BLAST score
                   521
                   7.0e-53
E value
                   127
Match length
                   80
% identity
                   (AC000098) Similar to Caenorhabditis unknown protein
NCBI Description
                   T03F1.1 (gb_U88169). [Arabidopsis thaliana]
                   5948
Seq. No.
                   5658 1.R1040
Contig ID
                   epx701105396.h1
5'-most EST
```

BLASTX Method NCBI GI g1762945 BLAST score 389

3.0e - 37E value 138 Match length 54 % identity

able to induce HR-like lesions [Nicotiana (U66269) ORF; NCBI Description

tabacum]

5949 Seq. No. 5658 2.R1040 Contig ID g5677292 5'-most EST BLASTX Method

q1762945 NCBI GI 209 BLAST score 5.0e-35 E value 107 Match length 70 % identity

able to induce HR-like lesions [Nicotiana (U66269) ORF; NCBI Description

tabacum]

Seq. No. 5950

5659 1.R1040 Contig ID

LIB3106-037-Q1-K1-C9 5'-most EST



Method BLASTX
NCBI GI g2499966
BLAST score 347
E value 2.0e-32
Match length 123
% identity 65

NCBI Description PHOTOSYSTEM I REACTION CENTRE SUBUNIT IV A PRECURSOR (PSI-E A) >gi_632722 bbs_151001 (S72356) photosystem I subunit

PSI-E [Nicotiana sylvestris, leaves, Peptide Chloroplast,

141 aa] [Nicotiana sylvestris]

Seq. No. 5951

Contig ID 5659 2.R1040

5'-most EST LIB3028-047-Q1-B1-A12

Seq. No. 5952

Contig ID 5666 1.R1040

5'-most EST LIB3 $\overline{0}$ 51-077-Q1-K1-F6

Method BLASTX
NCBI GI g1808656
BLAST score 1298
E value 1.0e-143
Match length 285
% identity 85

NCBI Description (Y10804) Ubiquitin activating enzyme E1 [Nicotiana tabacum]

Seq. No. 5953

Contig ID 5666_2.R1040 5'-most EST pmv700892693.h1

Method BLASTX
NCBI GI g1808656
BLAST score 654
E value 1.0e-112
Match length 247
% identity 80

NCBI Description (Y10804) Ubiquitin activating enzyme E1 [Nicotiana tabacum]

Seq. No. 5954

Contig ID 5669 1.R1040

5'-most EST LIB3170-043-Q1-J1-B10

Method BLASTX
NCBI GI g2739370
BLAST score 538
E value 6.0e-55
Match length 163
% identity 61

NCBI Description (AC002505) putative pectinesterase [Arabidopsis thaliana]

Seq. No. 5955

Contig ID 5671 1.R1040

5'-most EST LIB3065-022-Q1-N1-H5

Method BLASTX
NCBI GI g4539459
BLAST score 658
E value 2.0e-68
Match length 303
% identity 44



```
NCBI Description
                  (AL049500) putative protein [Arabidopsis thaliana]
Seq. No.
                  5956
                  5673 1.R1040
Contig ID
5'-most EST
                  g4286504
Method
                  BLASTX
NCBI GI
                  q2494222
BLAST score
                  344
                  5.0e-32
E value
Match length
                  83
                  72
% identity
NCBI Description
                  PROBABLE DYNEIN LIGHT CHAIN 1, CYTOPLASMIC >gi_470344
                  (U00043) No definition line found [Caenorhabditis elegans]
Seq. No.
                  5957
                  5673 3.R1040
Contig ID
                  jC-qmst02400073g05a1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2494222
BLAST score
                  240
E value
                  2.0e-20
Match length
                  60
                  70
% identity
NCBI Description PROBABLE DYNEIN LIGHT CHAIN 1, CYTOPLASMIC >gi 470344
                  (U00043) No definition line found [Caenorhabditis elegans]
                  5958
Seq. No.
Contig ID
                  5674 1.R1040
5'-most EST
                  LIB3107-021-Q1-K1-D6
Method
                  BLASTX
NCBI GI
                  q4432857
BLAST score
                  380
E value
                  3.0e-36
Match length
                  141
% identity
                  52
NCBI Description
                  (AC006300) hypothetical protein [Arabidopsis thaliana]
                  5959
Seq. No.
                  5674 2.R1040
Contig ID
5'-most EST
                  fde700876268.h1
                  5960
Seq. No.
                  5675 1.R1040
Contig ID
5'-most EST
                  jC-gmf102220080b08d1
                  5961
Seq. No.
                  5679 1.R1040
Contig ID
5'-most EST
                  uC-gmropic047h01b1
Method
                  BLASTX
NCBI GI
                  g4567247
BLAST score
                  727
E value
                  1.0e-76
Match length
                  335
```

Seq. No. 5962

% identity NCBI Description

(AC007070) unknown protein [Arabidopsis thaliana]



Contig ID 5679_3.R1040 5'-most EST 1eu701150389.h1

Seq. No. 5963

Contig ID 5680 1.R1040 5'-most EST ssr700554933.h1

Seq. No. 5964

Contig ID 5684 1.R1040

5'-most EST LIB3030-009-Q1-B1-H11

Method BLASTX
NCBI GI g2914700
BLAST score 786
E value 1.0e-83
Match length 277
% identity 61

NCBI Description (AC003974) tRNA-processing protein SEN3-like [Arabidopsis

thaliana]

Seq. No. 5965

Contig ID 5685 1.R1040

5'-most EST LIB3170-070-Q1-J1-H10

Seq. No. 5966

Contig ID 5686_1.R1040 5'-most EST txt700736063.h1

Method BLASTX
NCBI GI g118507
BLAST score 699
E value 2.0e-73
Match length 309
% identity 46

NCBI Description ALDEHYDE DEHYDROGENASE, DIMERIC NADP-PREFERRING (CLASS 3)

(TUMOR-ASSOCIATED ALDEHYDE DEHYDROGENASE) (HTC-ALDH) >gi_91936_pir__A30149 aldehyde dehydrogenase (NADP+) (EC 1.2.1.4) 3, tumor-associated - rat >gi_202833 (J03637)

aldehyde dehydrogenase [Rattus norvegicus]

Seq. No. 5967

Contig ID 5688_1.R1040

5'-most EST LIB3028-021-Q1-B1-H4

Seq. No. 5968

Contig ID 5696_1.R1040 5'-most EST leu701152149.h1

Method BLASTX
NCBI GI g2739044
BLAST score 1226
E value 1.0e-135
Match length 321
% identity 70

NCBI Description (AF024651) polyphosphoinositide binding protein Sshlp

[Glycine max]

Seq. No. 5969

Contig ID 5696 2.R1040 5'-most EST 95753550



Method BLASTX
NCBI GI g2739044
BLAST score 373
E value 1.0e-35
Match length 108
% identity 65

NCBI Description (AF024651) polyphosphoinositide binding protein Ssh1p

[Glycine max]

Seq. No. 5970

Contig ID 5697_1.R1040 5'-most EST vwf700678766.h1

Seq. No. 5971

Contig ID 5698 1.R1040

5'-most EST LIB3049-019-Q1-E1-A6

Method BLASTX
NCBI GI g1001607
BLAST score 290
E value 2.0e-25
Match length 136
% identity 39

NCBI Description (D64000) hypothetical protein [Synechocystis sp.]

Seq. No. 5972

Contig ID 5698 2.R1040

5'-most EST LIB3139-033-P1-N1-D8

Seq. No. 5973

Contig ID 5700 1.R1040

5'-most EST LIB3028-023-Q1-B1-E8

Seq. No. 5974

Contig ID 5704 1.R1040

5'-most EST LIB3056-009-Q1-N1-B6

Method BLASTX
NCBI GI g1916613
BLAST score 232
E value 8.0e-19
Match length 175
% identity 37

NCBI Description (U62029) acetyl-CoA carboxylase [Arabidopsis thaliana]

Seq. No. 5975

Contig ID 5710 1.R1040

5'-most EST LIB3028-025-Q1-B1-B11

Seq. No. 5976

Contig ID 5718_1.R1040

5'-most EST jC-gmf102220130d06d1

Seq. No. 5977

Contig ID 5729_1.R1040 5'-most EST fde700876539.h1

Method BLASTX NCBI GI g2271477 BLAST score 431



```
E value
                   2.0e-42
Match length
                   85
                   95
% identity
NCBI Description
                  (AF009631) AP47/50p [Arabidopsis thaliana]
Seq. No.
                   5978
Contig ID
                   5732 1.R1040
5'-most EST
                  LIB3170-042-Q1-J1-C10
Method
                   BLASTN
NCBI GI
                   g4263694
BLAST score
                   51
                   9.0e-20
E value
                   135
Match length
                   84
% identity
                  Arabidopsis thaliana chromosome II BAC F22D22 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
Seq. No.
                   5979
                   5732 2.R1040
Contig ID
5'-most EST
                   taw700658026.hl
Method
                   BLASTX
NCBI GI
                   q4263711
BLAST score
                   522
E value
                   3.0e-53
Match length
                   114
% identity
                   84
                   (AC006223) putative CCR4-associated transcription factor
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   5980
Contig ID
                   5735 1.R1040
                   wvk700683896.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   a1076579
BLAST score
                   676
E value
                   3.0e-71
Match length
                   167
% identity
                   75
                  alcohol dehydrogenase homolog ADH3a - tomato
NCBI Description
                   5981
Seq. No.
                   5736 1.R1040
Contig ID
5'-most EST
                   dpv701102630.hl
                   BLASTX
Method
                   g2245127
NCBI GI
                   166
BLAST score
E value
                   1.0e-11
```

Match length 74 % identity 47

NCBI Description (Z97344) hypothetical protein [Arabidopsis thaliana]

5982 Seq. No.

5740 1.R1040 Contig ID 5'-most EST vwf700676662.h1

Seq. No. 5983

5752 1.R1040 Contig ID



```
LIB3170-076-Q1-J1-A10
5'-most EST
                  5984
Seq. No.
                  5755 1.R1040
Contig ID
                  uC-qmflminsoy076f04b1
5'-most EST
                  BLASTX
Method
                  g4508068
NCBI GI
                  250
BLAST score
                  1.0e-20
E value
Match length
                  151
                  42
% identity
                 (AC005882) 3063 [Arabidopsis thaliana]
NCBI Description
                  5985
Seq. No.
                  5756 1.R1040
Contig ID
                   jC-gmf102220075f08d1
5'-most EST
                  BLASTN
Method
                  g1841474
NCBI GI
                   38
BLAST score
                   6.0e-12
E value
                   64
Match length
                   91
% identity
NCBI Description P.sativum mRNA for Myb-like protein (Myb26)
                   5986
Seq. No.
                   5763 1.R1040
Contig ID
                   LIB3028-039-Q1-B1-B5
5'-most EST
                   5987
Seq. No.
                   5763 2.R1040
Contig ID
                   LIB3107-080-Q1-K1-D1
5'-most EST
                   5988
Seq. No.
                   5763 3.R1040
Contig ID
                   uC-gmrominsoy115g03b1
5'-most EST
                   5989
Seq. No.
                   5777 1.R1040
Contig ID
                   LIB3028-040-Q1-B1-H4
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4490752
                   580
BLAST score
                   8.0e-60
E value
                   134
Match length
                   78
 % identity
                    (AL035708) putative serine/threonine protein kinase
NCBI Description
                    [Arabidopsis thaliana]
                   5990
 Seq. No.
                   5780 1.R1040
 Contig ID
                   uC-gmronoir015c09b1
 5'-most EST
                   BLASTX
 Method
                   g4512625
 NCBI GI
                   293
 BLAST score
```

- 300-

1131

5.0e-26

236

32

E value

Match length

% identity



NCBI Description (AC004793) ESTs gb T20423, gb AA712864, gb H76323 and gb Z25560 come from this gene. [Arabidopsis thaliana]

5991 Seq. No.

5784 1.R1040 Contig ID

5'-most EST LIB3028-041-Q1-B1-F12

Method BLASTX NCBI GI g3258637 BLAST score 537 E value 7.0e-55 Match length 145 % identity 70

NCBI Description (AF041050) 4-coumarate:CoA ligase [Populus tremuloides]

Seq. No.

5992 Contig ID 5786 1.R1040 5'-most EST zpv700763688.hl

Seq. No.

5993

Contig ID 5'-most EST 5787 1.R1040 trc700562171.h1

Method BLASTX NCBI GI q4544389 BLAST score 708 E value 2.0e-74 Match length 381 % identity

(AC007047) putative homeodomain protein [Arabidopsis NCBI Description

thaliana]

Seq. No.

5994

Contig ID 5791 1.R1040 5'-most EST zzp700832642.h1

Method BLASTX NCBI GI g1085002 BLAST score 552 E value 4.0e-56 Match length 295 % identity 42

NCBI Description mitochondrial carrier protein DIF-1 homolog -

Caenorhabditis elegans >gi_472902_emb_CAA53721_ (X76115)

carrier protein (c1) [Caenorhabditis elegans]

>gi_829102_emb_CAA88283_ (Z48240) DIF-1 [Caenorhabditis

elegans]

Seq. No.

5995

Contig ID

5791 2.R1040

5'-most EST

LIB3107-061-Q1-K1-G9

Method BLASTX NCBI GI q1085002 BLAST score 179 E value 6.0e-13 Match length 58 % identity 53

mitochondrial carrier protein DIF-1 homolog -NCBI Description

Caenorhabditis elegans >gi 472902 emb CAA53721 (X76115)

carrier protein (c1) [Caenorhabditis elegans]



>gi_829102_emb_CAA88283_ (Z48240) DIF-1 [Caenorhabditis
elegans]

Seq. No. 5996

Contig ID 5791_4.R1040 5'-most EST rlr700895682.h1

Seq. No. 5997

Contig ID 5791_5.R1040 5'-most EST bth700849349.h1

Seq. No. 5998

Contig ID 5792_1.R1040

5'-most EST LIB3028-042-Q1-B1-H12

Seq. No. 5999

Contig ID 5795_1.R1040 5'-most EST ssr700555019.h1

Seq. No. 6000

Contig ID 5797_1.R1040 5'-most EST hyd700726496.h1

Method BLASTX
NCBI GI g3831441
BLAST score 179
E value 3.0e-12
Match length 47
% identity 87

NCBI Description (AC005819) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6001

Contig ID 5797_2.R1040

5'-most EST jC-gmf102220073f08d1

Seq. No. 6002

Contig ID 5798_1.R1040
5'-most EST g4291359
Method BLASTX
NCBI GI g3335375
BLAST score 222
E value 1.0e-17
Match length 89

% identity 44
NCBI Description (AC003028) putative amidase [Arabidopsis thaliana]

Seq. No. 6003

Contig ID 5804_1.R1040

5'-most EST LIB3074-028-Q1-K1-F10

Method BLASTX
NCBI GI g4104816
BLAST score 505
E value 6.0e-51
Match length 136
% identity 70

NCBI Description (AF039662) ferredoxin-like protein [Capsicum annuum]

Seq. No. 6004

```
5804 2.R1040
Contig ID
5'-most EST
                  LIB3106-014-Q1-K1-B6
Method
                  BLASTX
NCBI GI
                  q4104816
BLAST score
                  473
                  2.0e-47
E value
Match length
                  130
% identity
                  (AF039662) ferredoxin-like protein [Capsicum annuum]
NCBI Description
Seq. No.
                  6005
                  5804 3.R1040
Contig ID
5'-most EST
                  LIB3106-092-Q1-K1-B12
Method
                  BLASTX
NCBI GI
                  g3023752
BLAST score
                  437
```

3.0e-43E value Match length 114 % identity 71

FERREDOXIN I PRECURSOR >gi_1418982_emb_CAA99756_ (Z75520) NCBI Description

ferredoxin-I [Lycopersicon esculentum]

6006 Seq. No.

5804 4.R1040 Contig ID

5'-most EST LIB3093-006-Q1-K1-A2

BLASTX Method NCBI GI g1834353 BLAST score 213 9.0e-17 E value Match length 67

% identity

NCBI Description (Y10986) hypothetical protein 194 [Arabidopsis thaliana]

6007 Seq. No.

5804 6.R1040 Contig ID

5'-most EST LIB3051-014-Q1-E1-F9

Method BLASTX g1834353 NCBI GI BLAST score 215 6.0e-17 E value Match length 245 29 % identity

(Y10986) hypothetical protein 194 [Arabidopsis thaliana] NCBI Description

Seq. No. 6008

Contig ID 5808 1.R1040

5'-most EST LIB3028-045-Q1-B1-D2

6009 Seq. No.

5809 1.R1040 Contig ID

LIB3092-037-Q1-K1-D11 5'-most EST

Method BLASTX NCBI GI g2829897 BLAST score 201 E value 3.0e-15 Match length 141 % identity 31



(AC002311) Unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

6010

Contig ID 5'-most EST 5809 2.R1040 uxk700670728.h1

Seq. No.

6011

Contig ID 5'-most EST 5809 3.R1040

LIB3107-063-Q1-K1-D6

Method NCBI GI BLAST score BLASTX q2829897 273

E value Match length 8.0e-24

% identity

91 66

NCBI Description

(AC002311) Unknown protein [Arabidopsis thaliana]

Seq. No.

6012

Contig ID

5811 1.R1040

5'-most EST

jC-gmst02400065e04a2

Method NCBI GI BLASTX g3288809 259

BLAST score E value Match length

9.0e-22 150

% identity NCBI Description

43 (AF013169) beta-glucosidase [Pyrococcus furiosus]

Seq. No.

6013

Contig ID

5812 1.R1040

5'-most EST

jC-qmle01810021e01d1

Method NCBI GI BLASTX q3142294

BLAST score E value

423 2.0e-41

Match length % identity

93 87

NCBI Description

(AC002411) Strong similarity to initiation factor eIF-2, gb U37354 from S. pombe. ESTs gb T41979, gb N37284 and gb N37529 come from this gene. [Arabidopsis thaliana]

Seq. No.

6014

Contig ID

5816 1.R1040

5'-most EST

jC-gmf102220056b06d1

Seq. No.

6015

Contig ID 5'-most EST

5818 1.R1040 ncj700980756.h1

Seq. No.

6016

Contig ID

5819 1.R1040

5'-most EST

LIB3028-047-Q1-B1-A9

Seq. No.

6017

Contig ID

5822 2.R1040

5'-most EST

LIB3039-052-Q1-E1-H8

Method

BLASTX



```
g3023637
NCBI GI
                   315
BLAST score
                   1.0e-28
E value
                   132
Match length
                   45
% identity
                   PROBABLE ATP-DEPENDENT RNA HELICASE HRH1 (DEAH BOX PROTEIN
NCBI Description
                   8) >gi_1362899_pir__A56236 probable RNA helicase 1 - human
                   >gi 80\overline{7}817 dbj BAA\overline{09}078 (D50487) RNA helicase (HRH1) [Homo
                   sapiens]
                   6018
Seq. No.
                   5822 3.R1040
Contig ID
                   LIB3094-062-Q1-K1-F2
5'-most EST
                    6019
Seq. No.
                    5827 1.R1040
Contig ID
                   g430\overline{4}284
5'-most EST
                    6020
Seq. No.
                    5829 1.R1040
Contiq ID
                    k117\overline{0}1203595.h2
5'-most EST
                    BLASTN
Method
                    g1553130
NCBI GI
BLAST score
                    153
                    2.0e-80
E value
                    321
Match length
                    87
% identity
                    Gossypium hirsutum ribosomal protein L44 isoform b (RL44),
NCBI Description
                    complete cds
                    6021
Seq. No.
                    5833 1.R1040
Contig ID
                    uxk700673268.h1
5'-most EST
                    BLASTX
Method
                    q4454032
NCBI GI
BLAST score
                    795
E value
                    1.0e-84
Match length
                    254
                    63
% identity
                    (AL035394) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    6022
                    5833 2.R1040
Contig ID
                    jC-gmst02400036b04d2
5'-most EST
                    BLASTX
Method
NCBI GI
                    g4454032
                    272
BLAST score
                    7.0e-24
E value
                    75
Match length
% identity
                    (AL035394) putative protein [Arabidopsis thaliana]
NCBI Description
                    6023
 Seq. No.
                    5833 3.R1040
Contig ID
```

LIB3170-068-Q1-K1-F2 5'-most EST

Method BLASTX NCBI GI g4454032



```
BLAST score
                   209
                   1.0e-16
E value
                   66
Match length
                   65
% identity
                  (AL035394) putative protein [Arabidopsis thaliana]
NCBI Description
                   6024
Seq. No.
                   5845 1.R1040
Contig ID
5'-most EST
                   LIB3028-051-Q1-B1-D6
                   6025
Seq. No.
                   5851 1.R1040
Contig ID
                   LIB3028-054-Q1-B1-G5
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4218120
                   368
BLAST score
                   2.0e-35
E value
                   83
Match length
% identity
                   (AL035353) Proline-rich APG-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   6026
Seq. No.
                   5851 2.R1040
Contig ID
5'-most EST
                   leu701155555.hl
                   BLASTX
Method
                   q4218120
NCBI GI
                   188
BLAST score
                   2.0e-14
E value
Match length
                   36
% identity
                   97
                    (AL035353) Proline-rich APG-like protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                    6027
                   5859 1.R1040
Contig ID
                    jC-qmro02910063g01a1
5'-most EST
                   BLASTX
Method
                   g3386621
NCBI GI
BLAST score
                    868
                    3.0e-93
E value
                    224
Match length
 % identity
                    (AC004665) unknown protein [Arabidopsis thaliana]
NCBI Description
 Seq. No.
                    6028
                    5859 2.R1040
 Contig ID
                    kmv7\overline{0}0737755.h1
 5'-most EST
                    BLASTX
```

Method g3386621 NCBI GI 700 BLAST score 6.0e-74 E value 156 Match length % identity

(AC004665) unknown protein [Arabidopsis thaliana] NCBI Description

6029 Seq. No.

1137



```
Contig ID 5859_3.R1040
5'-most EST LIB3092-052-Q1-K1-C11
Method BLASTX
```

NCBI GI g3386621 BLAST score 552 E value 1.0e-56 Match length 127 % identity 83

NCBI Description (AC004665) unknown protein [Arabidopsis thaliana]

Seq. No. 6030

5859 4.R1040 Contig ID 5'-most EST g4290314 Method BLASTX NCBI GI q3386621 BLAST score 536 E value 1.0e-54 Match length 162 % identity 67

NCBI Description (AC004665) unknown protein [Arabidopsis thaliana]

Seq. No. 6031

Contig ID 5865 1.R1040

5'-most EST LIB3029-002-Q1-B1-C5

Seq. No. 6032

Contig ID 5865_2.R1040 5'-most EST epx701105279.h1

Seq. No. 6033

Contig ID 5867 1.R1040

5'-most EST LIB3051-068-Q1-K1-G11

Method BLASTX
NCBI GI g1350707
BLAST score 212
E value 9.0e-17
Match length 52
% identity 73

NCBI Description 60S RIBOSOMAL PROTEIN L29 >gi 539923 pir JC2012 ribosomal

protein 17K - mouse >gi 404766 (L08651) ribosomal protein

[Mus musculus]

Seq. No. 6034

Contig ID 5867 2.R1040 5'-most EST g4395651 Method BLASTX NCBI GI q1350707 BLAST score 212 E value 6.0e-17 Match length 52 % identity 73

NCBI Description 60S RIBOSOMAL PROTEIN L29 >gi 539923 pir__JC2012 ribosomal

protein 17K - mouse >gi 404766 (L08651) ribosomal protein

[Mus musculus]

Seq. No. 6035

Contig ID 5869 1.R1040

NCBI GI

E value Match length

BLAST score

% identity

q18662

359 0.0e+00

374 99



```
5'-most EST
                  ssr700557855.hl
                  6036
Seq. No.
Contig ID
                  5874 1.R1040
5'-most EST
                  LIB3087-001-Q1-K1-D3
Method
                  BLASTX
NCBI GI
                  q2435511
BLAST score
                  337
E value
                  2.0e-31
Match length
                  111
% identity
                  57
                  (AF024504) contains similarity to prolyl 4-hydroxylase
NCBI Description
                  alpha subunit [Arabidopsis thaliana]
                  6037
Seq. No.
Contig ID
                  5877 1.R1040
                  jC-gmle01810087g09a1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q2924257
BLAST score
                  154
E value
                  2.0e-80
Match length
                  541
                  91
% identity
NCBI Description Tobacco chloroplast genome DNA
                  6038
Seq. No.
                  5877 2.R1040
Contig ID
                  awf700838734.hl
5'-most EST
Method
                  BLASTN
NCBI GI
                  g2924257
BLAST score
                  146
E value
                  4.0e-76
Match length
                  380
% identity
                  84
                 Tobacco chloroplast genome DNA
NCBI Description
                  6039
Seq. No.
Contig ID
                  5877 3.R1040
                  fde700872365.hl
5'-most EST
Method
                  BLASTN
NCBI GI
                  q1143165
BLAST score
                  196
E value
                  1.0e-106
Match length
                  430
% identity
                  86
                  Nicotiana tabacum ClpP protease (ClpP) mRNA, chloroplast
NCBI Description
                  gene encoding chloroplast protein, complete cds
                  6040
Seq. No.
                  5878 1.R1040
Contig ID
5'-most EST
                  epx701107225.hl
Method
                  BLASTN
```

1139



```
NCBI Description Glycine max hsp 70 gene
Seq. No.
                  6041
                  5893 1.R1040
Contig ID
5'-most EST
                  LIB3029-012-Q1-B1-D8
Seq. No.
                  6042
                  5895 1.R1040
Contig ID
5'-most EST
                  fC-gmst700837068d1
                  6043
Seq. No.
                  5899 1.R1040
Contig ID
5'-most EST
                  LIB3029-003-Q1-B1-H1
Method
                  BLASTX
NCBI GI
                  g4432827
BLAST score
                  247
                  9.0e-21
E value
                  176
Match length
% identity
                  36
NCBI Description (AC006593) putative ADP-ribose polymerase [Arabidopsis
                  thaliana]
Seq. No.
                  6044
                  5900 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy065f11b1
                  BLASTX
Method
NCBI GI
                  g1685005
BLAST score
                  296
E value
                  1.0e-26
Match length
                  121
                  46
% identity
NCBI Description (U32644) immediate-early salicylate-induced
                  glucosyltransferase [Nicotiana tabacum]
                  6045
Seq. No.
                  5906 1.R1040
Contig ID
5'-most EST
                  jC-gmst02400029h05a1
                  BLASTX
Method
NCBI GI
                  q1707412
BLAST score
                  884
E value
                  5.0e-95
Match length
                  417
% identity
                  44
NCBI Description (X95906) Cleavage and Polyadenylation Specifity Factor
                  protein [Bos taurus]
Seq. No.
                  6046
                  5917 1.R1040
Contig ID
5'-most EST
                  zsg701126164.h1
                  6047
Seq. No.
                  5917 2.R1040
Contig ID
5'-most EST
                  LIB3029-011-Q1-B1-H5
```

Contig ID 5925_1.R1040

5'-most EST LIB3029-012-Q1-B1-A6

BLAST score

E value

648 1.0e-67



```
6049
Seq. No.
                  5925 2.R1040
Contig ID
                  awf700836732.hl
5'-most EST
                  BLASTN
Method
                  g170005
NCBI GI
                  75
BLAST score
                   6.0e-34
E value
                   91
Match length
                   96
% identity
NCBI Description Soybean lectin (Le1) gene, complete cds
                   6050
Seq. No.
                   5927 1.R1040
Contig ID
                   uC-gmropic106c06b1
5'-most EST
                   BLASTN
Method
                   g1177602
NCBI GI
                   576
BLAST score
                   0.0e+00
E value
                   1016
Match length
                   89
% identity
NCBI Description P.sativum mRNA for pyruvate decarboxylase (PDC1)
                   6051
Seq. No.
                   5927_2.R1040
Contig ID
                   uC-gmropic088b10b1
5'-most EST
                   BLASTN
Method
                   g1177602
NCBI GI
                   322
BLAST score
                   0.0e+00
E value
                   693
Match length
% identity
NCBI Description P.sativum mRNA for pyruvate decarboxylase (PDC1)
                   6052
Seq. No.
                   5930 1.R1040
Contig ID
                   LIB3050-021-Q1-K1-B8
5'-most EST
                   BLASTX
Method
                   q3399767
NCBI GI
                   160
BLAST score
                   1.0e-10
E value
Match length
                   112
                   29
 % identity
                   (U76298) uclacyanin I [Arabidopsis thaliana] >gi_3831466
NCBI Description
                    (AC005700) uclacyanin I [Arabidopsis thaliana]
                    6053
 Seq. No.
                    5930 2.R1040
 Contig ID
                   LIB3109-009-Q1-K1-C1
 5'-most EST
                    6054
 Seq. No.
                    5933 1.R1040
 Contig ID
                   LIB3049-022-Q1-E1-G11
 5'-most EST
                    BLASTX
 Method
                    q1352681
 NCBI GI
```



Match length 174 % identity 72

NCBI Description

PROTEIN PHOSPHATASE 2C (PP2C) >gi_1076391_pir__ \$55457 phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis

thaliana >gi_633028_dbj_BAA07287_ (D38109) protein

phosphatase 2C [Arabidopsis thaliana]

Seq. No.

Contig ID

5933 2.R1040

5'-most EST LIB3170-059-Q1-J1-B3

6055

6056

Seq. No.

Contig ID 5'-most EST

5933_3.R1040 hrw701059606.h1

Method BLASTX
NCBI GI g1352681
BLAST score 783
E value 2.0e-83
Match length 285
% identity 57

NCBI Description PROTEIN PHOSPHATASE 2C (PP2C) >gi_1076391_pir__S55457 phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis

thaliana >gi_633028_dbj_BAA07287_ (D38109) protein

phosphatase 2C [Arabidopsis thaliana]

Seq. No. 6057

Contig ID

5933 4.R1040

5'-most EST

jC-gmro02910056a01a1

Method BLASTX
NCBI GI g1352681
BLAST score 362
E value 3.0e-34
Match length 165
% identity 48

NCBI Description

PROTEIN PHOSPHATASE 2C (PP2C) >gi 1076391 pir S55457 phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis

thaliana >gi 633028 dbj BAA07287 (D38109) protein

phosphatase 2C [Arabidopsis thaliana]

Seq. No. 6058

Contig ID 5936 1.R1040

5'-most EST LIB3039-010-Q1-E1-H9

Method BLASTX
NCBI GI g4097587
BLAST score 293
E value 3.0e-26
Match length 88
% identity 59

NCBI Description (U64926) NTGP5 [Nicotiana tabacum]

Seq. No.

6059

Contig ID 5936_2.R1040

5'-most EST uC-gmropic057f06b1

Seq. No.

6060

Contig ID 5936 3.R1040 5'-most EST jex700905711.h1



Seq. No. 6061

Contig ID 5937 1.R1040 5'-most EST leu701144592.h1

Method BLASTX NCBI GI q3687250 BLAST score 907 5.0e-98 E value Match length 208 80 % identity

(AC005169) putative arginine n-methyltransferase NCBI Description

[Arabidopsis thaliana]

6062 Seq. No.

Contig ID 5938 1.R1040

5'-most EST uC-qmflminsoy047q04b1

Method BLASTX NCBI GI g130720 BLAST score 289 E value 8.0e-26 112 Match length % identity 48

PROTEOLIPID PROTEIN PPA1 >qi 101508 pir A34633 probable NCBI Description

H+-transporting ATPase (EC 3.6.1.35) lipid-binding protein - yeast (Saccharomyces cerevisiae) >gi 172221 (M35294) proteolipid protein of proton ATPase [Saccharomyces

cerevisiae] >gi_500700 (U10399) Ppalp: Proteolipid protein

of proton ATPase [Saccharomyces cerevisiae]

Seq. No. 6063

5938 2.R1040 Contig ID 5'-most EST q5605673 Method BLASTX g130720 NCBI GI BLAST score 267 E value 4.0e-23 Match length 108

% identity 61

PROTEOLIPID PROTEIN PPA1 >gi_101508_pir__A34633 probable NCBI Description H+-transporting ATPase (EC 3.6.1.35) lipid-binding protein

- yeast (Saccharomyces cerevisiae) >gi_172221 (M35294) proteolipid protein of proton ATPase [Saccharomyces

cerevisiae] >gi 500700 (U10399) Ppalp: Proteolipid protein

of proton ATPase [Saccharomyces cerevisiae]

6064 Seq. No.

Contig ID 5951 1.R1040 sat701008670.h1 5'-most EST

Method BLASTX NCBI GI q1174448 BLAST score 599 E value 7.0e-62 Match length 222 % identity 54

NCBI Description TRANSLOCON-ASSOCIATED PROTEIN, ALPHA SUBUNIT PRECURSOR (TRAP-ALPHA) (SIGNAL SEQUENCE RECEPTOR ALPHA SUBUNIT)

(SSR-ALPHA) >gi 547391 (L32016) alpha-subunit; putative



[Arabidopsis thaliana]

```
      Seq. No.
      6065

      Contig ID
      5953_1.R1040

      5'-most EST
      LIB3049-053-Q1-E1-E7

      Method
      BLASTX
```

Method BLASTX
NCBI GI g3334323
BLAST score 958
E value 1.0e-104
Match length 193
% identity 92

NCBI Description GTP-BINDING PROTEIN SAR1A >gi_1314860 (U56929) Sar1 homolog

[Arabidopsis thaliana] >gi_2104532_gb_AAC78700.1

(AF001308) SAR1/GTP-binding secretory factor [Arabidopsis

thaliana] >gi 2104550 (AF001535) AGAA.4 [Arabidopsis

thaliana]

Seq. No. 6066

Contig ID 5956_1.R1040 5'-most EST txt700733748.h1

Method BLASTX
NCBI GI g4454471
BLAST score 287
E value 2.0e-25
Match length 75
% identity 73

NCBI Description (AC006234) putative G protein coupled receptor [Arabidopsis

thaliana]

Seq. No. 6067

Contig ID 5957 1.R1040

5'-most EST LIB3029-011-Q1-B1-B9

Seq. No. 6068

Contig ID 5958_1.R1040 5'-most EST zhf700952841.h1

Method BLASTX
NCBI GI g4115925
BLAST score 1023
E value 1.0e-111
Match length 298
% identity 69

NCBI Description (AF118222) contains similarity to RNA recognition motifs (Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana]

(Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana] >gi_4539439_emb_CAB40027.1_ (AL049523) RNA-binding protein

[Arabidopsis thaliana]

Seq. No. 6069

Contig ID 5958_2.R1040

5'-most EST LIB3049-039-Q1-E1-A8

Method BLASTX
NCBI GI g4115925
BLAST score 461
E value 7.0e-46
Match length 104
% identity 46

NCBI Description (AF118222) contains similarity to RNA recognition motifs



(Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana] >gi_4539439 emb_CAB40027.1_ (AL049523) RNA-binding protein [Arabidopsis thaliana]

Seq. No. 6070

Contig ID 5958 3.R1040 5'-most EST leu701151437.h1

Method BLASTX
NCBI GI g4115925
BLAST score 280
E value 2.0e-24
Match length 126
% identity 52

NCBI Description (AF118222) contains similarity to RNA recognition motifs

(Pfam: PF00076, Score=5.5e-23, N=2) [Arabidopsis thaliana] >gi 4539439 emb CAB40027.1_ (AL049523) RNA-binding protein

[Arabidopsis thaliana]

Seq. No. 6071

Contig ID 5958_4.R1040 5'-most EST vzy700755441.h1

Seq. No. 6072

Contig ID 5958_5.R1040 5'-most EST gsv701048154.h1

Seq. No. 6073

Contig ID 5960_1.R1040 5'-most EST ncj700984222.h1

Method BLASTX
NCBI GI g118926
BLAST score 709
E value 7.0e-75
Match length 201
% identity 67

NCBI Description DESSICATION-RELATED PROTEIN CLONE PCC13-62 PRECURSOR

>gi_320600_pir__E45509 desiccation-related protein (clone PCCI3-62) - Craterostigma plantagineum >gi_167479 (M62991) dessication-related protein [Craterostigma plantagineum] >gi_227781 prf_ 1710351E abscisic acid responsive protein E

[Craterostigma plantagineum]

Seq. No. 6074

Contig ID 5966_1.R1040 5'-most EST smc700750004.h1

Method BLASTX
NCBI GI g542190
BLAST score 613
E value 2.0e-63
Match length 287
% identity 45

NCBI Description hypothetical protein 1087 - maize >gi 459269 emb CAA54960

(X78029) transcribed sequence 1087 [Zea mays]

Seq. No. 6075

Contig ID 5966 2.R1040 5'-most EST smc700745314.h1



Method BLASTX
NCBI GI g542190
BLAST score 373
E value 8.0e-36
Match length 133
% identity 44

NCBI Description hypothetical protein 1087 - maize >gi_459269_emb_CAA54960_

(X78029) transcribed sequence 1087 [Zea mays]

Seq. No. 6076

Contig ID 5967 1.R1040

5'-most EST LIB3139-032-P1-N1-E1

Method BLASTX
NCBI GI 94567229
BLAST score 336
E value 2.0e-31
Match length 138
% identity 46

NCBI Description (AC007119) putative pectin methylesterase [Arabidopsis

thaliana]

Seq. No.

Contig ID 5968 1.R1040

5'-most EST LIB3106-113-Q1-K1-D12

6077

Method BLASTX
NCBI GI g2935416
BLAST score 1293
E value 1.0e-143
Match length 302
% identity 81

NCBI Description (AF047896) isoflavone reductase homolog [Betula pendula]

Seq. No.

6078

Contig ID 5970_1.R1040

5'-most EST LIB3029-011-Q1-B1-B8

Seq. No. 6079

Contig ID 5979_1.R1040

5'-most EST jC-qmst02400042e08a1

Method BLASTX
NCBI GI g3914431
BLAST score 1144
E value 1.0e-125
Match length 249
% identity 86

NCBI Description PROTEASOME COMPONENT C8 (MACROPAIN SUBUNIT C8)

(MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT C8)

>gi 2285802 dbj BAA21651 (D78173) 26S proteasome alpha

subunit [Spinacia oleracea]

Seq. No. 6080

Contig ID 5979 2.R1040 5'-most EST kl1701208578.h1

Method BLASTX
NCBI GI g3914431
BLAST score 631
E value 6.0e-66



Match length 151 % identity 87

PROTEASOME COMPONENT C8 (MACROPAIN SUBUNIT C8) NCBI Description (MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT C8)

>gi_2285802_dbj_BAA21651_ (D78173) 26S proteasome alpha

subunit [Spinacia oleracea]

6081 Seq. No.

Contig ID 5983 1.R1040 5'-most EST leu701146005.hl

Method BLASTX NCBI GI q2660676 BLAST score 811 E value 1.0e-86 206 Match length 73 % identity

(AC002342) Dreg-2 like protein [Arabidopsis thaliana] NCBI Description

Seq. No.

6082 5985 1.R1040 Contig ID 5'-most EST awf700840391.h1

Method BLASTX NCBI GI q81886 BLAST score 753 E value 6.0e-80 Match length 205 % identity 71

NCBI Description tonoplast intrinsic protein alpha - kidney bean

6083 Seq. No.

Contig ID 5987 1.R1040 5'-most EST zhf700952328.h1

Seq. No. 6084

Contig ID 5987 2.R1040

5'-most EST jC-gmf102220114f01a1

Seq. No. 6085

5994 1.R1040 Contig ID crh700853990.h1 5'-most EST

BLASTX Method NCBI GI g2245066 BLAST score 1398 E value 1.0e-155 Match length 460 % identity 58

NCBI Description (Z97342) Beta-Amylase [Arabidopsis thaliana]

Seq. No. 6086

5994 2.R1040 Contig ID 5'-most EST rca700996317.hl

Method BLASTX NCBI GI g2245066 BLAST score 259 E value 2.0e-22 Match length 128 % identity 45



(Z97342) Beta-Amylase [Arabidopsis thaliana] NCBI Description

Seq. No.

6087

Contig ID 5'-most EST 5995 1.R1040 kl1701212350.h1

Seq. No.

6088

Contig ID 5'-most EST 5996 1.R1040 uaw700662796.hl

Seq. No.

6089 5996 2.R1040 Contig ID

5'-most EST

zzp700833912.hl BLASTX

Method NCBI GI BLAST score E value

g2459431 254 1.0e-21

Match length % identity

158 44

NCBI Description

(AC002332) unknown protein [Arabidopsis thaliana]

Seq. No.

Contig ID

5997 1.R1040 hrw701058407.h1

5'-most EST Method NCBI GI

BLASTX q3395431

BLAST score E value

1.0e-13 107 Match length 34 % identity

NCBI Description

(AC004683) unknown protein [Arabidopsis thaliana]

Seq. No. Contig ID 6091

186

5999 1.R1040

5'-most EST

LIB3029-010-Q1-B1-C3

Seq. No.

6092

Contig ID 5'-most EST 5999 4.R1040 asn701135609.hl

Seq. No.

6093

Contig ID

6000 1.R1040

g3080451

1116

5'-most EST

jC-gmro02910009d06a1 BLASTX

Method NCBI GI BLAST score E value

1.0e-125 Match length 362 % identity 62

NCBI Description

(AL022605) hypothetical protein [Arabidopsis thaliana]

Seq. No.

6094

Contig ID

6000 2.R1040

5'-most EST

LIB3107-066-Q1-K1-A3

Method NCBI GI BLAST score BLASTX g3080451 385



E value 3.0e-37 Match length 141 % identity 54

(AL022605) hypothetical protein [Arabidopsis thaliana] NCBI Description

#1. · · ·

6095 Seq. No.

6007 1.R1040 Contig ID

5'-most EST LIB3029-010-Q1-B1-B7

BLASTX Method NCBI GI g2652938 BLAST score 902 E value 1.0e-97 Match length 190 % identity 87

NCBI Description (Z47554) orf [Zea mays]

Seq. No.

6096 Contig ID 6011 1.R1040 5'-most EST uaw700661589.h1

Method BLASTX NCBI GI g126078 BLAST score 510 E value 1.0e-51 Match length 184 62 % identity

LATE EMBRYOGENESIS ABUNDANT PROTEIN D-34 (LEA D-34) NCBI Description

>gi_81554_pir__S04046 embryonic abundant protein gD-34 upland cotton >gi_18501_emb_CAA31594_ (X13206) D-34 Lea protein [Gossypium hirsutum] >gi 167385 (M19389) storage protein [Gossypium hirsutum] >gi_226556_prf__1601521F Lea

D-34 gene [Saguinus oedipus]

6097 Seq. No.

Contig ID 6013 1.R1040

5'-most EST LIB3029-010-Q1-B1-A5

6098 Seq. No.

Contig ID 6016 1.R1040

5'-most EST LIB3029-010-Q1-B1-A8

Method BLASTX NCBI GI q4510383 BLAST score 559 E value 2.0e-57 Match length 190 % identity

(AC007017) unknown protein [Arabidopsis thaliana] NCBI Description

6099 Seq. No.

6018 1.R1040 Contig ID

5'-most EST LIB3051-010-Q1-E1-G10

Method BLASTX NCBI GI q3080439 BLAST score 151 E value 9.0e-10 Match length 75 % identity 44

(AL022605) putative protein [Arabidopsis thaliana] NCBI Description



Seq. No. 6100

Contig ID 6018_2.R1040 5'-most EST kl1701207641.h1

Seq. No. 6101

Contig ID 6022_1.R1040 5'-most EST wvk700685854.h1

Seq. No. 6102

Contig ID 6025 1.R1040

5'-most EST LIB3106-032-Q1-K1-B12

Method BLASTN
NCBI GI g303856
BLAST score 160
E value 2.0e-84
Match length 384
% identity 85

NCBI Description Rice mRNA for ubiquitin protein fused to a ribosomal

protein, complete cds

Seq. No. 6103

Contig ID 6029 1.R1040

5'-most EST LIB3107-071-Q1-K1-B8

Seq. No. 6104

Contig ID 6029 2.R1040

5'-most EST LIB3051-016-Q1-E1-G10

Method BLASTX
NCBI GI g1435021
BLAST score 374
E value 1.0e-35
Match length 168
% identity 53

NCBI Description (D26575) DNA-binding protein [Daucus carota]

Seq. No. 6105

Contig ID 6030_1.R1040

5'-most EST LIB3051-112-Q1-K1-G6

Method BLASTX
NCBI GI g3152596
BLAST score 629
E value 2.0e-65
Match length 309
% identity 50

NCBI Description (AC002986) YUP8H12R.36 [Arabidopsis thaliana]

Seq. No. 6106

Contig ID 6030_3.R1040

5'-most EST fC-gmle700682314f1

Seq. No. 6107

Contig ID 6033_1.R1040 5'-most EST g5126510 Method BLASTX NCBI GI g3776578 BLAST score 332



E value 1.0e-30 Match length 152 % identity 42

NCBI Description (AC005388) ESTs gb_F13915 and gb_F13916 come from this

gene. [Arabidopsis thaliana]

Seq. No. 6108

Contig ID 6035_1.R1040 5'-most EST crh700855751.h1

Method BLASTX
NCBI GI g3878119
BLAST score 202
E value 3.0e-15
Match length 92
% identity 49

NCBI Description (Z49068) similar to GTP-binding protein; cDNA EST

EMBL:M89111 comes from this gene; cDNA EST EMBL:D27709 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST

yk353

Seq. No. 6109

Contig ID 6040 1.R1040

5'-most EST LIB3029-009-Q1-B1-C4

Method BLASTN
NCBI GI g2347097
BLAST score 67
E value 4.0e-29
Match length 171
% identity 85

NCBI Description Arabidopsis thaliana ubiquitin-specific protease (AtUBP3)

mRNA, complete cds

Seq. No. 6110

Contig ID 6041_1.R1040 5'-most EST pxt700943230.h1

Seq. No. 6111

Contig ID 6043_1.R1040 5'-most EST zzp700833930.h1

Method BLASTX
NCBI GI 93402711
BLAST score 455
E value 2.0e-83
Match length 239
% identity 43

NCBI Description (AC004261) putative RNA-binding protein [Arabidopsis

thaliana]

Seq. No. 6112

Contig ID 6043_2.R1040

5'-most EST LIB3051-094-Q1-K1-C7

Method BLASTN
NCBI GI g2347085
BLAST score 92
E value 6.0e-44
Match length 156



```
% identity
                   90
                  Phaseolus vulgaris putative osmoprotector PvLEA-18
NCBI Description
                   (Pvlea-18) mRNA, complete cds
                   6113
Seq. No.
                   6048 1.R1040
Contig ID
                  leu701156292.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g4337179
BLAST score
                  314
E value
                  9.0e-29
                  145
Match length
                   49
% identity
                   (AC006416) This gene is continued on the 5' end of BAC
NCBI Description
                  T12M14. [Arabidopsis thaliana]
                   6114
Seq. No.
Contig ID
                   6049 1.R1040
5'-most EST
                   jC-gmf102220108f01a1
Method
                  BLASTX
NCBI GI
                  g2618705
BLAST score
                  876
E value
                   2.0e-94
Match length
                  271
% identity
                   68
                   (AC002510) putative ABC transporter, 5' partial
NCBI Description
                   [Arabidopsis thaliana]
                   6115
Seq. No.
                   6072 1.R1040
Contig ID
5'-most EST
                   gsv701046930.h1
Method
                  BLASTX
NCBI GI
                  q2342727
BLAST score
                  281
                   5.0e-25
E value
Match length
                  68
                  71
% identity
NCBI Description
                  (AC002341) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   6116
                   6075 1.R1040
Contig ID
5'-most EST
                  has700548234.h1
Method
                  BLASTX
NCBI GI
                   g2956717
BLAST score
                   237
E value
                  8.0e-20
Match length
                  125
% identity
                   46
                  (Y09722) beta-carotene hydroxylase 2 [Capsicum annuum]
NCBI Description
Seq. No.
                   6117
```

Contig ID 6078 1.R1040

5'-most EST LIB3029-008-Q1-B1-G2

Seq. No.

6118

Contig ID 6080_1.R1040 5'-most EST epx701106595.h1



Method BLASTX
NCBI GI g2842757
BLAST score 192
E value 3.0e-14
Match length 109
% identity 43

NCBI Description LIGHT-INDUCIBLE PROTEIN CPRF-2 >gi_1806261_emb_CAA41453_ (X58577) DNA-binding protein; bZIP type [Petroselinum

crispum]

Seq. No. 6119

Contig ID 6080_2.R1040

5'-most EST jC-gmf102220102e09a1

Method BLASTX
NCBI GI g2842757
BLAST score 245
E value 2.0e-21
Match length 156
% identity 46

NCBI Description LIGHT-INDUCIBLE PROTEIN CPRF-2 >gi 1806261 emb CAA41453

(X58577) DNA-binding protein; bZIP type [Petroselinum

crispum]

Seq. No. 6120

Contig ID 6080 3.R1040

5'-most EST uC-gmropic113g07b1

Method BLASTX
NCBI GI g100163
BLAST score 230
E value 6.0e-19
Match length 147
% identity 40

NCBI Description light-induced protein CPRF-2 - parsley

Seq. No. 6121

Contig ID 6082_1.R1040 5'-most EST pcp700990468.h1

Method BLASTX
NCBI GI g1914683
BLAST score 905
E value 1.0e-97
Match length 245
% identity 74

NCBI Description (Y12013) RAD23, isoform I [Daucus carota]

Seq. No. 6122

Contig ID 6082 2.R1040

5'-most EST LIB3093-001-Q1-K1-G12

Method BLASTX
NCBI GI g1914683
BLAST score 413
E value 3.0e-40
Match length 107
% identity 78

NCBI Description (Y12013) RAD23, isoform I [Daucus carota]

Seq. No. 6123



Contig ID 6082_3.R1040 5'-most EST crh700852624.h1

Method BLASTX
NCBI GI g1914683
BLAST score 289
E value 7.0e-26
Match length 74

% identity 76

NCBI Description (Y12013) RAD23, isoform I [Daucus carota]

Seq. No. 6124

Contig ID 6087 1.R1040

5'-most EST LIB3029-008-Q1-B1-D3

Method BLASTX
NCBI GI g130718
BLAST score 338
E value 1.0e-31
Match length 100
% identity 61

NCBI Description ACID PHOSPHATASE PRECURSOR 1 >gi_170370 (M83211) acid

phosphatase type 1 [Lycopersicon esculentum] >gi_170372 (M67474) acid phosphatase type 5 [Lycopersicon esculentum] >gi_445121 prf 1908427A acid phosphatase 1 [Lycopersicon

esculentum]

Seq. No. 6125

Contig ID 6089 1.R1040

5'-most EST LIB3087-008-Q1-K1-A6

Seq. No. 6126

Contig ID 6089 2.R1040

5'-most EST LIB3139-013-P1-N1-G7

Seq. No. 6127

Contig ID 6095_1.R1040

5'-most EST jC-gmfl02220108h05a1

Method BLASTX
NCBI GI g4325342
BLAST score 548
E value 5.0e-56
Match length 240
% identity 53

NCBI Description (AF128393) No definition line found [Arabidopsis thaliana]

Seq. No. 6128

Contig ID 6102 1.R1040

5'-most EST LIB3170-064-Q1-J1-E9

Seq. No. 6129

Contig ID 6102 2.R1040

5'-most EST jC-qmf102220057h09d1

Method BLASTX
NCBI GI g1171866
BLAST score 273
E value 6.0e-24
Match length 52
% identity 94



NADH-UBIQUINONE OXIDOREDUCTASE 20 KD SUBUNIT PRECURSOR NCBI Description (COMPLEX I-20KD) (CI-20KD) >gi_629601_pir__S48826 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST - wild cabbage >gi 562282 emb CAA57725 (X82274) PSST subunit of NADH: ubiquinone oxidoreductase [Brassica oleracea] 6130 Seq. No. 6106 1.R1040 Contig ID 5'-most EST pmv700893391.h1 Method BLASTN NCBI GI g1161251 BLAST score 61 2.0e-25 E value Match length 253 % identity 81 NCBI Description Glycine max nucleosome assembly protein 1 (SNAP-1) mRNA, complete cds Seq. No. 6131 Contig ID 6112 1.R1040 gsv701055056.h1 5'-most EST BLASTX Method NCBI GI q3261634 BLAST score 187 E value 1.0e-13 190 Match length % identity NCBI Description (279700) hypothetical protein Rv0976c [Mycobacterium tuberculosis] Seq. No. 6132 6114 1.R1040 Contig ID 5'-most EST LIB3029-008-Q1-B1-D1 Method BLASTX NCBI GI q3914002 BLAST score 923 E value 1.0e-100 Match length 189 % identity 93 MITOCHONDRIAL LON PROTEASE HOMOLOG 1 PRECURSOR >gi 2935279 NCBI Description (AF033862) Lon protease [Arabidopsis thaliana] 6133 Seq. No. 6115 1.R1040 Contig ID 5'-most EST LIB3107-016-Q1-K1-D3 Method BLASTX

NCBI GI q4371285 BLAST score 171 E value 4.0e-12 Match length 63 % identity

(AC006260) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 6134

6116 1.R1040 Contig ID 5'-most EST epx701108611.h1

1155



```
6135
Seq. No.
Contig ID
                   6118 1.R1040
5'-most EST
                   LIB3\overline{1}39-103-P1-N1-A9
Seq. No.
                   6136
                   6118 2.R1040
Contig ID
5'-most EST
                   q5058185
                   6137
Seq. No.
                   6120 1.R1040
Contig ID
5'-most EST
                   LIB3029-007-Q1-B1-H7
                   6138
Seq. No.
Contig ID
                   6121 1.R1040
5'-most EST
                   LIB3029-008-Q1-B1-A11
                   6139
Seq. No.
                   6126 1.R1040
Contig ID
```

5'-most EST zzp700832478.hl Method BLASTX NCBI GI q3482967 BLAST score 758 E value 1.0e-149

Match length 359 % identity 72

(AL031369) Protein phosphatase 2C-like protein [Arabidopsis NCBI Description thaliana] >gi 4559345 gb AAD23006.1 AC006585 1 (AC006585)

protein phosphatase 2C [Arabidopsis thaliana]

6140 Seq. No. 6127 1.R1040 Contig ID 5'-most EST LIB3167-042-P1-K1-H4 Method BLASTX q2078350 NCBI GI BLAST score 721

E value 2.0e-76 Match length 199 % identity 70

(U95923) transaldolase [Solanum tuberosum] NCBI Description

Seq. No. 6141 6127 2.R1040 Contig ID 5'-most EST leu701152348.h1 Method BLASTX NCBI GI g2078350

BLAST score 529 E value 6.0e-54 Match length 136 71 % identity

(U95923) transaldolase [Solanum tuberosum] NCBI Description

Seq. No. 6142

6129 1.R1040 Contig ID

5'-most EST LIB3029-008-Q1-B1-B2

Seq. No. 6143

Contig ID 6133 1.R1040



5'-most EST sat701006002.h2

Method BLASTX
NCBI GI g2244939
BLAST score 346
E value 6.0e-32
Match length 304
% identity 31

NCBI Description (297339) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6144

Contig ID 6133 2.R1040

5'-most EST jC-gmle01810044h11d1

Seq. No. 6145

Contig ID 6133_3.R1040 5'-most EST vzy700756777.h1

Seq. No. 6146

Contig ID 6136 1.R1040 5'-most EST asn701133071.h1

Method BLASTN
NCBI GI g18729
BLAST score 64
E value 2.0e-27
Match length 68
% identity 99

NCBI Description Soybean (Glycine max) 18S ribosomal RNA

Seq. No. 6147

Contig ID 6136_2.R1040 5'-most EST sat701008527.h1

Method BLASTN
NCBI GI g18729
BLAST score 269
E value 1.0e-149
Match length 458
% identity 90

NCBI Description Soybean (Glycine max) 18S ribosomal RNA

Seq. No. 6148

Contig ID 6144_1.R1040 5'-most EST leu701157711.h1

Method BLASTX
NCBI GI g2623297
BLAST score 143
E value 1.0e-08
Match length 49
% identity 55

NCBI Description (AC002409) unknown protein [Arabidopsis thaliana]

>gi_3790583 (AF079180) RING-H2 finger protein RHC1a

[Arabidopsis thaliana]

Seq. No. 6149

Contig ID 6144 2.R1040

5'-most EST jC-gmfl02220063c05a1

Seq. No. 6150



```
Contig ID
                   6145 1.R1040
5'-most EST
                  LIB3167-027-P1-K1-A6
Method
                  BLASTX
                  q2462834
NCBI GI
BLAST score
                  364
                  2.0e-34
E value
Match length
                  182
                  43
% identity
                  (AF000657) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  6151
Seq. No.
Contig ID
                  6145 3.R1040
5'-most EST
                  wvk700684271.hl
Method
                  BLASTX
NCBI GI
                  q3337357
BLAST score
                  165
E value
                  5.0e-13
Match length
                  68
```

% identity 62 NCBI Description (AC004481) hypothetical protein [Arabidopsis thaliana]

6152 Seq. No. Contig ID 6145 4.R1040 jex700906373.hl 5'-most EST Method BLASTX q3337357 NCBI GI

BLAST score 278 E value 5.0e-25 Match length 86 % identity 65

NCBI Description (AC004481) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6153

Contig ID 6150 1.R1040 5'-most EST zsg701121114.hl

Seq. No. 6154

Contig ID 6158 1.R1040

5'-most EST LIB3029-007-Q1-B1-E9

Seq. No. 6155

6159 1.R1040 Contig ID

5'-most EST LIB3029-007-Q1-B1-D3

Method BLASTX NCBI GI g2204234 BLAST score 485 E value 4.0e-49 Match length 118 % identity

(Y13862) enoyl-ACP reductase [Nicotiana tabacum] NCBI Description

6156 Seq. No.

6160 1.R1040 Contig ID uaw700661722.hl 5'-most EST

BLASTX Method NCBI GI q3264830 BLAST score 212



E value 7.0e-17 Match length 149 % identity 44

NCBI Description (AF072405) cotton fiber expressed protein 2 [Gossypium

hirsutum]

Seq. No. 6157

Contig ID 6169 1.R1040

5'-most EST LIB3029-007-Q1-B1-C7

Method BLASTX
NCBI GI g2497540
BLAST score 412
E value 2.0e-51
Match length 161
% identity 66

NCBI Description PYRUVATE KINASE, CHLOROPLAST ISOZYME G

Seq. No. 6158

Contig ID 6171_1.R1040

5'-most EST LIB3109-015-Q1-K1-F1

Seq. No. 6159

Contig ID 6171 2.R1040

5'-most EST LIB3106-023-Q1-K1-C3

Method BLASTX
NCBI GI g3769673
BLAST score 224
E value 3.0e-18
Match length 120
% identity 48

NCBI Description (AF095285) Tic20 [Pisum sativum]

Seq. No. 6160

Contig ID 6171_4.R1040 5'-most EST jsh701069881.h1

Seq. No. 6161

Contig ID 6172_1.R1040

5'-most EST LIB3029-007-Q1-B1-D1

Seq. No. 6162

Contig ID 6182 1.R1040

5'-most EST LIB3029-007-Q1-B1-A1

Method BLASTX
NCBI GI g2982283
BLAST score 369
E value 3.0e-35
Match length 126
% identity 57

NCBI Description (AF051226) PREG-like protein [Picea mariana]

Seq. No. 6163

Contig ID 6195 1.R1040

5'-most EST LIB3040-023-Q1-E1-E6

Method BLASTX NCBI GI g4160280 BLAST score 1414



```
E value 1.0e-157

Match length 302
% identity 82

NCBI Description (AJ006224) purple acid phosphatase [Ipomoea batatas]

Seq. No. 6164
Contig ID 6196 1.R1040
```

5'-most EST vzy700754476.h1

Method BLASTX

NCBI GI g2407613

BLAST score 705

E value 8.0e-74

Match length 442

% identity 40

NCBI Description (AF017995) 3-phosphoinositide dependent protein kinase-1 [Homo sapiens] >gi_2505936_emb_CAA75341_ (Y15056) PkB kinase [Homo sapiens] >gi_4505695_ref_NP_002604.1_pPDPK1_

3-phosphoinositide dependent protein kinase-1

 Seq. No.
 6165

 Contig ID
 6201_1.R1040

 5'-most EST
 rlr700899564.h1

 Method
 BLASTN

 NCBI GI
 g4519414

 NAST grant
 182

NCBI GI 9451941 BLAST score 182 E value 1.0e-97 Match length 470 % identity 85

NCBI Description Citrus unshiu (CitVATP c-2) mRNA for vacuolar H+-ATPase c

subunit, complete cds

Seq. No. 6166

Contig ID 6203 1.R1040

5'-most EST LIB3029-006-Q1-B1-H1

Method BLASTX
NCBI GI g2829918
BLAST score 194
E value 8.0e-15
Match length 71
% identity 54

NCBI Description (AC002291) similar to "tub" protein gp_U82468_2072162

[Arabidopsis thaliana]

Seq. No. 6167

Contig ID 6208_1.R1040 5'-most EST kl1701203850.h1

Seq. No. 6168

Contig ID 6217 1.R1040

5'-most EST LIB3029-006-Q1-B1-E9

Seq. No. 6169

Contig ID 6221_1.R1040 5'-most EST hyd700726494.h1

Method BLASTX
NCBI GI g3212116
BLAST score 277



E value 4.0e-24 Match length 127 % identity 44

NCBI Description (Y17393) prefoldin subunit 2 [Mus musculus]

Seq. No. 6170

Contig ID 6221 2.R1040

5'-most EST LIB3051-096-Q1-K1-A8

Method BLASTX
NCBI GI g3212116
BLAST score 260
E value 2.0e-22
Match length 102
% identity 47

NCBI Description (Y17393) prefoldin subunit 2 [Mus musculus]

Seq. No. 6171

Contig ID 6221_6.R1040

5'-most EST LIB3049-038-Q1-E1-E10

Seq. No. 6172

Contig ID 6222 1.R1040

5'-most EST LIB3029-006-Q1-B1-B9

Method BLASTX
NCBI GI g1946300
BLAST score 164
E value 2.0e-11
Match length 36
% identity 46

NCBI Description (Y12529) hypothetical protein [Silene latifolia]

Seq. No. 6173

Contig ID 6225 1.R1040

5'-most EST LIB3029-006-Q1-B1-C12

Seq. No. 6174

Contig ID 6229 1.R1040

5'-most EST LIB3049-024-Q1-E1-A4

Seq. No. 6175

Contig ID 6230_1.R1040

5'-most EST jC-gmro02800033g12a1

Method BLASTX
NCBI GI g3334349
BLAST score 510
E value 1.0e-51
Match length 176
% identity 72

NCBI Description GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE) (GLYRS)

>gi_2564215_emb_CAA05162_ (AJ002062) glycyl-tRNA synthetase

[Arabidopsis thaliana]

Seq. No. 6176

Contig ID 6230 2.R1040

5'-most EST LIB3039-012-Q1-E1-A2

Method BLASTX NCBI GI q3334349



BLAST score 1181 E value 1.0e-130 Match length 343 68 % identity

NCBI Description GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE) (GLYRS)

>gi_2564215_emb_CAA05162_ (AJ002062) glycyl-tRNA synthetase

[Arabidopsis thaliana]

Seq. No.

6177 Contig ID 6230 4.R1040 5'-most EST kl1701214567.hl

Method BLASTX NCBI GI q3334349 BLAST score 609 E value 6.0e-69 Match length 175 78 % identity

GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE) (GLYRS) NCBI Description

>gi 2564215 emb CAA05162 (AJ002062) glycyl-tRNA synthetase

[Arabidopsis thaliana]

6178 Seq. No.

6230 5.R1040 Contig ID 5'-most EST pmv700890116.hl Method BLASTX

NCBI GI q3334349 BLAST score 449 E value 6.0e-45 Match length 101 % identity 86

NCBI Description GLYCYL-TRNA SYNTHETASE (GLYCINE--TRNA LIGASE) (GLYRS)

>gi_2564215_emb_CAA05162_ (AJ002062) glycyl-tRNA synthetase

[Arabidopsis thaliana]

Seq. No. 6179

Contig ID 6231 1.R1040

5'-most EST LIB3029-006-Q1-B1-D11

6180

Seq. No.

6238 1.R1040 Contig ID

5'-most EST uC-gmronoir018a07b1

BLASTN Method NCBI GI g3241920 BLAST score 38 E value 2.0e-11 Match length 338 % identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MAEl, complete sequence [Arabidopsis thaliana]

Seq. No. 6181

6238 2.R1040 Contig ID

 $LIB3\overline{0}51-084-Q1-K1-C8$ 5'-most EST

6182 Seq. No.

Contig ID 6244 1.R1040 5'-most EST gsv701046253.hl



```
Method BLASTN
NCBI GI g436031
BLAST score 112
E value 9.0e-56
Match length 328
% identity 84
```

NCBI Description Nicotiana tabacum (TSC40-4) 60S ribosomal protein L34 mRNA,

complete cds

Seq. No. 6183

Contig ID 6244_2.R1040

5'-most EST LIB3039-035-Q1-E1-G6

Method BLASTX
NCBI GI 9730558
BLAST score 471
E value 5.0e-47
Match length 95
% identity 95

NCBI Description 60S RIBOSOMAL PROTEIN L34 >gi 1076636 pir S48027 ribosomal

protein L34 - common tobacco >gi_2129964_pir__S48028 ribosomal protein L34.e, cytosolic - common tobacco >gi_436030 (L27089) 60S ribosomal protein L34 [Nicotiana tabacum] >gi_436032 (L27107) 60S ribosomal protein L34

[Nicotiana tabacum]

Seq. No. 6184

Contig ID 6244_3.R1040 5'-most EST leu701149904.h1

Method BLASTX
NCBI GI 9730557
BLAST score 379
E value 2.0e-36
Match length 75
% identity 95

NCBI Description 60S RIBOSOMAL PROTEIN L34 >gi_2119150_pir__S60476 ribosomal

protein L34 - garden pea >gi 498908 (U10047) ribosomal

protein L34 homolog [Pisum sativum]

Seq. No. 6185

Contig ID 6247 1.R1040

5'-most EST jC-gmle01810005c06a1

Method BLASTX
NCBI GI g3183454
BLAST score 648
E value 2.0e-67
Match length 278
% identity 47

NCBI Description HYPOTHETICAL 30.7 KD PROTEIN IN MCPC-KINA INTERGENIC REGION

>gi_2632217_emb_CAA10859_ (AJ222587) YkwC protein [Bacillus
subtilis] >gi_2633767_emb_CAB13269_ (Z99111) similar to
3-hydroxyisobutyrate dehydrogenase [Bacillus subtilis]

Seq. No. 6186

Contig ID 6247 2.R1040 5'-most EST ncj700986168.h1

Method BLASTX NCBI GI g3183454



```
BLAST score
                      211
E value
                      1.0e-16
Match length
                      87
                      49
% identity
                     HYPOTHETICAL 30.7 KD PROTEIN IN MCPC-KINA INTERGENIC REGION
NCBI Description
                      >gi_2632217_emb_CAA10859_ (AJ222587) YkwC protein [Bacillus
subtilis] >gi_2633767_emb_CAB13269_ (Z99111) similar to
                      3-hydroxyisobutyrate dehydrogenase [Bacillus subtilis]
Seq. No.
                      6187
Contig ID
                      6251 1.R1040
```

5'-most EST uC-gmropic114f01b1 Seq. No. 6188

Contig ID 6258_1.R1040
5'-most EST txt700732131.h1
Method BLASTX
NCBI GI g3176098

NCBI GI g3176098 BLAST score 1128 E value 1.0e-123 Match length 312 % identity 71

NCBI Description (Y15036) annexin [Medicago truncatula]

 Seq. No.
 6189

 Contig ID
 6261_1.R1040

 5'-most EST
 LIB3107-010-Q1-K1-F3

 Method
 BLASTX

 NCBI GI
 g4376815

NCBI GI g4376818
BLAST score 400
E value 1.0e-38
Match length 178
% identity 45

NCBI Description (AE001637) GutQ/KpsF Family Sugar-P Isomerase [Chlamydia

pneumoniae]

Seq. No. 6190 Contin ID 6266

Contig ID 6266 1.R1040 5'-most EST zhf700962155.h1

Method BLASTX
NCBI GI g2911068
BLAST score 744
E value 6.0e-79
Match length 145
% identity 88

NCBI Description (AL021960) G10 - like protein [Arabidopsis thaliana]

Seq. No. 6191

Contig ID 6284_1.R1040 5'-most EST ncj700978444.h1

Method BLASTX
NCBI GI 94240305
BLAST score 275
E value 8.0e-24
Match length 215
% identity 33

NCBI Description (AB020715) KIAA0908 protein [Homo sapiens]



Seq. No. 6192 Contig ID 6285

6285_1.R1040

5'-most EST

LIB3029-005-Q1-B1-D10

Method BLASTX
NCBI GI g2244816
BLAST score 201
E value 1.0e-15
Match length 60
% identity 75

NCBI Description (Z97336) hypothetical protein [Arabidopsis thaliana]

Seq. No.

6193

Contig ID

6290 1.R1040

5'-most EST

LIB3094-059-Q1-K1-C1

Method BLASTX
NCBI GI 9729775
BLAST score 482
E value 1.0e-48
Match length 116
% identity 78

NCBI Description HEAT SHOCK FACTOR PROTEIN HSF8 (HEAT SHOCK TRANSCRIPTION

FACTOR 8) (HSTF 8) (HEAT STRESS TRANSCRIPTION FACTOR)
>gi_100264_pir__S25481 heat shock transcription factor 8 Peruvian tomato >gi 19492 emb CAA47869 (X67600) heat shock

transcription factor 8 [Lycopersicon peruvianum]

Seq. No. 6194

Contig ID 6291_1.R1040

5'-most EST LIB3049-023-Q1-E1-H9

Seq. No. 6195

Contig ID 6293_1.R1040 5'-most EST kl1701205910.h1

Method BLASTX
NCBI GI g3915020
BLAST score 2619
E value 0.0e+00
Match length 639
% identity 85

NCBI Description SUCROSE-PHOSPHATE SYNTHASE (UDP-GLUCOSE-FRUCTOSE-PHOSPHATE

GLUCOSYLTRANSFERASE) >gi_1022365_emb_CAA91217_ (Z56278)

sucrose phosphate synthase [Vicia faba]

Seq. No. 6196

Contig ID 6295 1.R1040 5'-most EST 94287190 Method BLASTX NCBI GI 93434971 BLAST score 278 E value 2.0e-24 Match length 73

% identity 73

NCBI Description (AB008105) ethylene responsive element binding factor 3

[Arabidopsis thaliana]

Seq. No. 6197



Contig ID 6295_2.R1040 5'-most EST uxk700672508.h1

Method BLASTX
NCBI GI g1208496
BLAST score 243
E value 1.0e-20
Match length 56
% identity 82

NCBI Description (D38124) EREBP-3 [Nicotiana tabacum]

Seq. No. 6198

Contig ID 6301 1.R1040

5'-most EST LIB3051-116-Q1-K1-G3

Method BLASTX
NCBI GI g3935184
BLAST score 578
E value 2.0e-59
Match length 197
% identity 56

NCBI Description (AC004557) F17L21.27 [Arabidopsis thaliana]

Seq. No. 6199

Contig ID 6307_1.R1040 5'-most EST 6HA-02-Q1-B1-F1

Method BLASTX
NCBI GI g558606
BLAST score 382
E value 6.0e-37
Match length 109
% identity 62

NCBI Description (X70688) Nthsp18p [Nicotiana tabacum]

Seq. No. 6200

Contig ID 6313_1.R1040 5'-most EST pmv700892451.h1

Method BLASTN
NCBI GI g2997590
BLAST score 523
E value 0.0e+00
Match length 1032
% identity 88

NCBI Description Pisum sativum glucose-6-phosphate/phosphate-translocator

precursor (GPT) mRNA, nuclear gene encoding plastid

protein, complete cds

Seq. No. 6201

Contig ID 6313_2.R1040 5'-most EST jex700908125.h1

Method BLASTN
NCBI GI g2997590
BLAST score 156
E value 2.0e-82
Match length 329
% identity 91

NCBI Description Pisum sativum glucose-6-phosphate/phosphate-translocator

precursor (GPT) mRNA, nuclear gene encoding plastid

protein, complete cds

1166

...



6202 Seq. No. 6313 3.R1040 Contig ID wvk700683529.h1 5'-most EST 6203 Seq. No. Contig ID 6315 1.R1040 LIB3049-012-Q1-E1-F2 5'-most EST Method BLASTX q3913633

NCBI GI BLAST score 540 E value 5.0e-55 Match length 162 % identity 64

HYPOTHETICAL PROTEIN F8A5.25 >qi 2462742 (AC002292) Unknown NCBI Description protein [Arabidopsis thaliana]

Seq. No. 6204

6316 1.R1040 Contig ID 5'-most EST hyd700725490.hl

BLASTX Method NCBI GI g4006859 BLAST score 319 3.0e-29 E value 180 Match length 46 % identity

NCBI Description (Z99707) putative protein [Arabidopsis thaliana]

6205 Seq. No.

6319 1.R1040 Contig ID

5'-most EST LIB3049-006-Q1-E1-A2

6206 Seq. No.

Contig ID 6322 1.R1040

5'-most EST LIB3029-005-Q1-B1-A2

Method BLASTX NCBI GI g1620898 BLAST score 994 1.0e-108 E value 265 Match length 73 % identity

NCBI Description (D87957) protein involved in sexual development [Homo

sapiens]

6207 Seq. No.

Contig ID 6323 1.R1040 5'-most EST vzy700752174.hl

6208 Seq. No.

Contig ID 6324 1.R1040

5'-most EST LIB3029-004-Q1-B1-G7

Method BLASTX NCBI GI q2407068 BLAST score 224 E value 9.0e-18 Match length 122 39 % identity



NCBI Description (AF014955) TFAR19 [Homo sapiens]

Seq. No. 6209

Contig ID 6326 1.R1040

5'-most EST LIB3029-004-Q1-B1-F11

Seg. No. 6210

Contig ID 6331_1.R1040

5'-most EST LIB3072-060-Q1-K1-D7

Method BLASTN
NCBI GI g18653
BLAST score 508
E value 0.0e+00
Match length 548
% identity 98

NCBI Description Soybean gene for heat shock protein Gmhsp18.5-C (class I)

Seq. No. 6211

6331 2.R1040 Contig ID 5'-most EST g5606297 Method BLASTN NCBI GI g18653 BLAST score 290 E value 1.0e-162 579 Match length 89 % identity

NCBI Description Soybean gene for heat shock protein Gmhsp18.5-C (class I)

Seq. No. 6212

Contig ID 6331 3.R1040

5'-most EST LIB3170-041-Q1-J1-A9

Method BLASTN
NCBI GI g18655
BLAST score 631
E value 0.0e+00
Match length 813
% identity 95

NCBI Description Soybean heat-shock gene hs6871 sequence

Seq. No. 6213

Contig ID 6331_4.R1040 5'-most EST epx701105644.h1

Method BLASTN
NCBI GI g18655
BLAST score 251
E value 1.0e-139
Match length 387
% identity 93

NCBI Description Soybean heat-shock gene hs6871 sequence

Seq. No. 6214

Contig ID 6331_5.R1040 5'-most EST epx701107217.h1

Method BLASTN
NCBI GI g18655
BLAST score 396
E value 0.0e+00



Match length 396 % identity 100

NCBI Description Soybean heat-shock gene hs6871 sequence

Seq. No.

6215

Contig ID

6331_6.R1040

5'-most EST

LIB3072-019-Q1-E1-H5

Method NCBI GI BLAST score E value

g18653 205 1.0e-111

BLASTN

Match length % identity

313 91

NCBI Description

Soybean gene for heat shock protein Gmhsp18.5-C (class I)

Seq. No.

6216

Contig ID 5'-most EST

6331_7.R1040 hrw701060690.h1

Method NCBI GI BLAST score BLASTN g169984 351

E value
Match length
% identity

0.0e+00 359 99

NCBI Description

Soybean (Glycine max) low MW heat shock protein gene

(Gmhsp17.6-L)

Seq. No.

6217

Contig ID 5'-most EST

6331_8.R1040 epx701104850.h1

Method NCBI GI BLAST score E value BLASTN g1431738 147 4.0e-77

Match length % identity

235 91

NCBI Description

Soybean (Glycine max) low MW heat shock protein gene

(Gmhsp17.5-M)

Seq. No.

6218

Contig ID 5'-most EST

6331_10.R1040 hrw701061389.h1

Method NCBI GI BLASTN g1431738 182

BLAST score E value Match length

6.0e-98 266

93

% identity
NCBI Description

Soybean (Glycine max) low MW heat shock protein gene

(Gmhsp17.5-M)

Seq. No.

6219

Contig ID

6331 12.R1040

5'-most EST

uC-gmrominsoy134f10b1

Method NCBI GI BLASTN g18653

BLAST score

63



E value 5.0e-27 Match length 191 % identity 91

NCBI Description Soybean gene for heat shock protein Gmhsp18.5-C (class I)

Seq. No. 6220

Contig ID 6331_13.R1040 5'-most EST hrw701062361.h1

Method BLASTN
NCBI GI g18653
BLAST score 214
E value 1.0e-117
Match length 262
% identity 96

NCBI Description Soybean gene for heat shock protein Gmhsp18.5-C (class I)

Seq. No. 6221

Contig ID 6331_14.R1040 5'-most EST zsg701128782.h1

Method BLASTN
NCBI GI g169986
BLAST score 87
E value 2.0e-41
Match length 154
% identity 90

NCBI Description Soybean (Glycine max) heat shock protein (Gmhsp17.5-E)

gene, complete cds

Seq. No. 6222

Contig ID 6340_1.R1040 5'-most EST gsv701050470.h1

Seq. No. 6223

Contig ID 6340_3.R1040

5'-most EST uC-gmropic111b05b1

Seq. No. 6224

Contig ID 6346_1.R1040

5'-most EST LIB3049-010-Q1-E1-C11

Method BLASTX
NCBI GI 94539303
BLAST score 469
E value 9.0e-47
Match length 164
% identity 56

NCBI Description (AL049480) putative protein [Arabidopsis thaliana]

Seq. No. 6225

Contig ID 6346_2.R1040 5'-most EST dpv701100045.h2

Method BLASTX
NCBI GI 94539303
BLAST score 245
E value 9.0e-21
Match length 101
% identity 51

NCBI Description (AL049480) putative protein [Arabidopsis thaliana]

% identity

88

```
6226
Seq. No.
                  6348 1.R1040
Contig ID
5'-most EST
                  LIB3029-004-Q1-B1-E5
Method
                  BLASTX
NCBI GI
                  g2213590
BLAST score
                  216
                  3.0e-17
E value
Match length
                  128
                  38
% identity
NCBI Description
                  (AC000348) T7N9.10 [Arabidopsis thaliana]
Seq. No.
                   6227
Contig ID
                  6350 1.R1040
5'-most EST
                  LIB3072-024-Q1-E1-G1
                  BLASTN
Method
NCBI GI
                  g460802
BLAST score
                  54
E value
                  3.0e-21
Match length
                  238
% identity
                  81
NCBI Description P.amygdalus, Batsch (Texas) ole1 mRNA
                   6228
Seq. No.
Contig ID
                   6350 2.R1040
5'-most EST
                  LIB3072-056-Q1-K1-D10
Method
                  BLASTX
NCBI GI
                  g3914199
BLAST score
                  222
E value
                  5.0e-18
Match length
                  141
% identity
                   40
                  OLEOSIN 1 >gi_460803_emb_CAA55008_ (X78118) oleosin [Prunus
NCBI Description
                  dulcis]
                   6229
Seq. No.
                   6354 1.R1040
Contig ID
5'-most EST
                  g5606027
                  BLASTX
Method
NCBI GI
                  g3688188
BLAST score
                   2538
                  0.0e+00
E value
Match length
                   608
                   80
% identity
                   (AL031804) pyruvate decarboxylase-1 (Pdc1) [Arabidopsis
NCBI Description
                   thaliana]
                   6230
Seq. No.
                   6354 3.R1040
Contig ID
5'-most EST
                  txt700734851.hl
Method
                  BLASTN
NCBI GI
                  g1177604
BLAST score
                   165
E value
                   1.0e-87
Match length
                   429
```

NCBI Description P.sativum mRNA for pyruvate decarboxylase (PDC2)



```
Seq. No. 6231
```

Contig ID 6354_4.R1040 5'-most EST bnu700967544.h1

Method BLASTX
NCBI GI g3688188
BLAST score 306
E value 6.0e-28
Match length 94
% identity 66

NCBI Description (AL031804) pyruvate decarboxylase-1 (Pdc1) [Arabidopsis

thaliana]

Seq. No. 6232

Contig ID 6359 1.R1040

5'-most EST uC-gmflminsoy055c05b1

Method BLASTX
NCBI GI g4006829
BLAST score 707
E value 2.0e-74
Match length 207
% identity 67

NCBI Description (AC005970) putative protein kinase [Arabidopsis thaliana]

Seq. No. 6233

Contig ID 6359_2.R1040 5'-most EST kl1701211035.h1

Method BLASTX
NCBI GI g3805765
BLAST score 161
E value 8.0e-11
Match length 48
% identity 67

NCBI Description (AC005693) putative protein kinase [Arabidopsis thaliana]

Seq. No. 6234

Contig ID 6359_3.R1040 5'-most EST ssr700553651.h1

Seq. No. 6235

Contig ID 6361_1.R1040 5'-most EST uaw700663807.h1

Method BLASTX
NCBI GI g452519
BLAST score 181
E value 1.0e-12
Match length 131
% identity 20

NCBI Description (D26362) similar to Human homolog of Drosophila female

sterile homeotic mRNA (HUMFSHG) [Homo sapiens]

Seq. No. 6236

Contig ID 6362_1.R1040 5'-most EST txt700731402.h1

Method BLASTN NCBI GI g18638 BLAST score 318



```
E value 1.0e-179
Match length 384
% identity 98
```

NCBI Description Soybean Gy3 gene for glycinin subunit G3

 Seq. No.
 6237

 Contig ID
 6364_1.R1040

 5'-most EST
 zzp700832050.h1

 Method
 BLASTX

 MEDICAL
 21733440

Method BLASTX
NCBI GI g1723440
BLAST score 495
E value 2.0e-49
Match length 246
% identity 43

NCBI Description HYPOTHETICAL 35.9 KD PROTEIN C56F8.08 IN CHROMOSOME I

>gi_1204230_emb_CAA93579_ (Z69728) unknown

[Schizosaccharomyces pombe]

Seq. No. 6238

Contig ID 6364_2.R1040

5'-most EST LIB3051-053-Q1-K2-C4

Seq. No. 6239

Contig ID 6368 1.R1040

5'-most EST uC-gmrominsoy188h04b1

Method BLASTX
NCBI GI g2827552
BLAST score 331
E value 1.0e-30
Match length 96
% identity 61

NCBI Description (AL021635) predicted protein [Arabidopsis thaliana]

Seq. No. 6240

Contig ID 6373 1.R1040 5'-most EST jex700906521.h1

Method BLASTN
NCBI GI g18731
BLAST score 1060
E value 0.0e+00
Match length 1621
% identity 97

NCBI Description Soybean RPB1-B1 gene for the largest subunit of RNA

polymerase II (EC 2.7.7.6)

Seq. No. 6241

6385 1.R1040 Contig ID 5'-most EST q5509589 Method BLASTX g1063276 NCBI GI BLAST score 787 E value 5.0e - 84Match length 220 69 % identity

NCBI Description (X92893) geranylgeranyl pyrophosphate synthase

[Catharanthus roseus]



Seq. No. 6242 Contig ID 6386

Contig ID 6386_1.R1040 5'-most EST zpv700758409.h1

Method BLASTX
NCBI GI g633890
BLAST score 973
E value 1.0e-105
Match length 280
% identity 69

NCBI Description (S72926) glucose and ribitol dehydrogenase homolog [Hordeum

vulgare]

Seq. No. 6243

Contig ID 6386_4.R1040 5'-most EST uxk700668322.h1

Method BLASTX
NCBI GI g633890
BLAST score 170
E value 5.0e-12
Match length 44
% identity 73

NCBI Description (S72926) glucose and ribitol dehydrogenase homolog [Hordeum

vulgare]

Seq. No. 6244

Contig ID 6390 1.R1040

5'-most EST LIB3039-010-Q1-E1-C8

Method BLASTX
NCBI GI g3617770
BLAST score 552
E value 1.0e-56
Match length 138
% identity 73

NCBI Description (Y14329) threonyl-tRNA synthetase [Arabidopsis thaliana]

Seq. No. 6245

Contig ID 6390 2.R1040

5'-most EST LIB3029-003-Q1-B1-G10

Method BLASTX
NCBI GI g2501064
BLAST score 302
E value 2.0e-27
Match length 113
% identity 58

NCBI Description PROBABLE THREONYL-TRNA SYNTHETASE, CYTOPLASMIC

(THREONINE--TRNA LIGASE) (THRRS) >gi_2191162 (AF007270) Similar to threonyl-tRNA synthetase; coded for by A.

thaliana cDNA R65376 [Arabidopsis thaliana]

Seq. No. 6246

Contig ID 6397 1.R1040

5'-most EST LIB3072-011-Q1-E1-D1

Method BLASTN
NCBI GI g1141781
BLAST score 233
E value 1.0e-128
Match length 317

% identity 93
NCBI Description Vigna radiata EM protein mRNA, complete cds

Seq. No. 6247

Contig ID 6399_1.R1040 5'-most EST txt700734368.h1

Seq. No. 6248

Contig ID 6399 2.R1040 5'-most EST crh700853112.h1

Seq. No. 6249

Contig ID 6400 1.R1040

5'-most EST LIB3029-003-Q1-B1-E11

Method BLASTX
NCBI GI g3738230
BLAST score 286
E value 1.0e-25
Match length 85
% identity 60

NCBI Description (AB007790) DREB2A [Arabidopsis thaliana]

>gi 4126706_dbj_BAA36705 (AB016570) DREB2A [Arabidopsis

thaliana]

Seq. No. 6250

Contig ID 6401 1.R1040

5'-most EST uC-gmrominsoy047c11b1

Method BLASTX
NCBI GI g1001449
BLAST score 154
E value 6.0e-10
Match length 91

Match length 91 % identity 41

NCBI Description (D63999) hypothetical protein [Synechocystis sp.]

Seq. No. 6251

Contig ID 6404 1.R1040

5'-most EST LIB3139-022-P1-N1-H3

Method BLASTX
NCBI GI g3551958
BLAST score 714
E value 2.0e-75
Match length 197
% identity 68

NCBI Description (AF082032) senescence-associated protein 12 [Hemerocallis

hybrid cultivar]

Seq. No. 6252

Contig ID 6404 2.R1040

5'-most EST uC-gmrominsoy207h01b1

Method BLASTX
NCBI GI g2708750
BLAST score 501
E value 2.0e-50
Match length 253
% identity 46

NCBI Description (AC003952) putative physical impedence protein [Arabidopsis



thaliana]

Seq. No.

6253

Contig ID

6405 1.R1040

5'-most EST LIB3094-061-Q1-K1-B3

Seq. No.

6254

Contig ID

6405 2.R1040

5'-most EST

LIB3049-009-Q1-E1-A11

Method NCBI GI BLASTX q4325372

BLAST score E value

665 3.0e-69

Match length % identity

282 54

NCBI Description

(AF128396) contains similarity to protein disulfide

isomerases [Arabidopsis thaliana]

Seq. No.

6255

Contig ID 5'-most EST

6405 3.R1040 uxk700669734.hl

Seq. No.

6256

Contig ID

6405 4.R1040

5'-most EST

uC-gmronoir058g07b1

Seq. No.

6257

Contig ID 5'-most EST

6405 5.R1040 awf700836446.hl

Seq. No.

6258

Contig ID

6405 6.R1040

5'-most EST

LIB3106-035-Q1-K1-D1

Seq. No.

6259

Contig ID

6405 8.R1040

5'-most EST

jC-gmf102220139g01a1

Seq. No.

6260

Contig ID

6406 1.R1040

5'-most EST

LIB3050-006-Q1-E1-G2

Seq. No.

6261

Contig ID

6408 1.R1040

5'-most EST

LIB3049-012-Q1-E1-H8

Seq. No.

6262

Contig ID

6408 2.R1040

5'-most EST

LIB3107-039-Q1-K1-D10

Seq. No.

6263

Contig ID

6411 1.R1040

5'-most EST

LIB3170-057-Q1-K1-D6

Method NCBI GI BLAST score BLASTX q2129944

410 1.0e-39

E value



193 Match length % identity 47

RNA-binding protein RZ-1 - wood tobacco NCBI Description

>gi_1395193_dbj_BAA12064_ (D83696) RNA-binding protein RZ-1
[Nicotiana sylvestris] >gi_1435062_dbj_BAA06012_ (D28861)

RNA binding protein, RZ-1 [Nicotiana sylvestris]

Seq. No.

6264

Contig ID 6416 1.R1040

5'-most EST LIB3029-003-Q1-B1-C4

Seq. No.

6265

Contig ID

6428 1.R1040

5'-most EST

uC-gmflminsoy056c09b1

Method BLASTX NCBI GI g4049341 BLAST score 166 3.0e-11 E value

Match length 146 % identity 35

(AL034567) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No.

6266

Contig ID 5'-most EST

6433 1.R1040 ncj700987169.h1

Method BLASTX NCBI GI g169363 BLAST score 377 9.0e-36 E value Match length 151 54

% identity

NCBI Description (M75856) PVPR3 [Phaseolus vulgaris]

Seq. No.

6267

Contig ID

6433 2.R1040

5'-most EST

LIB3087-001-Q1-K1-D6

Seq. No.

6268

Contig ID

6434 1.R1040

5'-most EST

LIB3107-021-Q1-K1-E3

Seq. No.

6269

Contig ID

6434 2.R1040

5'-most EST

LIB3051-002-Q1-E1-H11

Seq. No.

6270

Contig ID

6435 1.R1040 gsv701054428.h1

5'-most EST

6271

Seq. No. Contig ID

6440 1.R1040

5'-most EST

LIB3074-035-Q1-K1-C9

Method NCBI GI BLASTX g3377820

BLAST score E value

436 5.0e-43

Match length

123



% identity 70

NCBI Description (AF076275) contains similarity to coatomer zeta chains

[Arabidopsis thaliana]

Seq. No. 6272

Contig ID 6440 2.R1040

5'-most EST LIB3040-023-Q1-E1-D9

Method BLASTX
NCBI GI g3377820
BLAST score 280
E value 5.0e-25
Match length 64
% identity 80

NCBI Description (AF076275) contains similarity to coatomer zeta chains

[Arabidopsis thaliana]

Seq. No. 6273

6441 1.R1040 Contig ID 5'-most EST $q429\overline{2}967$ Method BLASTN NCBI GI g438898 BLAST score 136 E value 3.0e-70 Match length 407 85 % identity

NCBI Description Soybean glutamine phosphoribosylpyrophosphate

amidotransferase mRNA, complete cds

Seq. No. 6274

Contig ID 6442_1.R1040 5'-most EST trc700561456.h1

Seq. No. 6275

Contig ID 6442 2.R1040

5'-most EST LIB3072-015-Q1-E1-D10

Seq. No. 6276

Contig ID 6445_1.R1040 5'-most EST gsv701046109.h1

Seq. No. 6277

Contig ID 6445_2.R1040 5'-most EST uC-gmropic038e03b1

Seq. No. 6278

Contig ID 6452_1.R1040 5'-most EST uaw700666570.h1

Method BLASTN
NCBI GI g534940
BLAST score 57
E value 4.0e-23
Match length 78
% identity 93

NCBI Description S.tuberosum mitochondrial DNA for ribosomal protein S10

Seq. No. 6279

Contig ID 6453 1.R1040



5'-most EST ncj700981278.h1

Method BLASTN
NCBI GI g169930
BLAST score 392
E value 0.0e+00
Match length 897
% identity 87

NCBI Description Glycine max calcium dependent protein kinase mRNA

Seq. No. 6280

Contig ID 6453_2.R1040 5'-most EST fC-gmse700674513a2

Method BLASTX
NCBI GI g116054
BLAST score 309
E value 1.0e-28
Match length 76
% identity 74

NCBI Description CALCIUM-DEPENDENT PROTEIN KINASE SK5 (CDPK)

>gi_280393_pir__A43713 calcium-dependent protein kinase (EC
2.7.1.-) - soybean >gi_169931 (M64987) Glycine max calcium

dependent protein kinase mRNA. [Glycine max]

Seq. No. 6281

Contig ID 6456_1.R1040 5'-most EST leu701153778.h1

Method BLASTX
NCBI GI g115515
BLAST score 402
E value 8.0e-39
Match length 149
% identity 37

NCBI Description CALMODULIN >gi_71683_pir_MCAA calmodulin - alfalfa

>gi_19579_emb_CAA366 $\overline{4}4$ _(X52398) calmodulin (AA 1-149) [Medicago sativa] >gi_170070 (L01430) calmodulin [Glycine

max] >gi 170074 (L01432) calmodulin [Glycine max]

>gi 506852 (L20507) calmodulin [Vigna radiata] >gi 4103957

(AF030032) calmodulin [Phaseolus vulgaris]

>gi_1583767_prf__2121384A calmodulin [Glycine max]
>gi_1583769_prf__2121384C calmodulin [Glycine max]

Seq. No. 6282

Contig ID 6456 3.R1040

5'-most EST LIB3029-002-Q1-B1-E4

Seq. No. 6283

Contig ID 6456_4.R1040 5'-most EST jsh701069327.h1

Method BLASTX
NCBI GI g4379369
BLAST score 331
E value 7.0e-31
Match length 122
% identity 33

NCBI Description (X90560) Calmodulin [Cloning vector pBluescript]

Seq. No. 6284



```
6456 5.R1040
Contig ID
5'-most EST
                  LIB3106-007-Q1-K1-A9
                  6285
Seq. No.
Contig ID
                  6457 1.R1040
                  sat701009851.h1
5'-most EST
                   6286
Seq. No.
Contig ID
                  6457 2.R1040
5'-most EST
                  fC-qmse700675662f2
Method
                  BLASTX
NCBI GI
                  q2760839
BLAST score
                  177
E value
                  2.0e-12
                  62
Match length
                  29
% identity
NCBI Description (AC003105) putative receptor kinase [Arabidopsis thaliana]
                  6287
Seq. No.
                  6457 3.R1040
Contig ID
5'-most EST
                  crh700850651.h1
Seq. No.
                  6288
Contig ID
                  6461 1.R1040
5'-most EST
                  uC-gmropic066d11b1
Seq. No.
                  6289
Contig ID
                  6464 1.R1040
5'-most EST
                  LIB3029-002-Q1-B1-F5
Seq. No.
                  6290
Contig ID
                  6465 1.R1040
5'-most EST
                  g4396524
Method
                  BLASTX
NCBI GI
                  q4091806
BLAST score
                  824
E value
                  4.0e-88
Match length
                  309
% identity
                  60
                  (AF052585) CONSTANS-like protein 2 [Malus domestica]
NCBI Description
Seq. No.
                  6291
                  6467 1.R1040
Contig ID
5'-most EST
                  LIB3029-002-Q1-B1-G1
Method
                  BLASTX
NCBI GI
                  q3036807
BLAST score
                  513
E value
                  6.0e-52
Match length
                  167
% identity
                  62
NCBI Description
                  (AL022373) putative protein [Arabidopsis thaliana]
```

Seq. No. 6292 6478 1.R1040 Contig ID 5'-most EST leu701145927.h1

Method BLASTX NCBI GI g1084415



BLAST score 398 E value 1.0e-38 Match length 123 % identity 65

NCBI Description RNA-binding protein - Wood tobacco >gi_624925_dbj_BAA05170_

(D26182) RNA-binding glycine rich protein (RGP-2)

[Nicotiana sylvestris]

Seq. No. Contig ID 6293 6478 2.R1040

5'-most EST

jC-gmst02400028g05a1

Method BLASTX
NCBI GI g2493318
BLAST score 453
E value 6.0e-45
Match length 102
% identity 83

NCBI Description BLUE COPPER PROTEIN PRECURSOR >gi_562779_emb_CAA80963_

(Z25471) blue copper protein [Pisum sativum]

>gi_1098264_prf__2115352A blue Cu protein [Pisum sativum]

Seq. No.

Contig ID

6294 6478 3.R1040

5'-most EST

uC-gmrominsoy316g06b1

Method BLASTX
NCBI GI g1084415
BLAST score 344
E value 2.0e-32
Match length 110
% identity 65

NCBI Description RNA-binding protein - Wood tobacco >gi 624925 dbj BAA05170

(D26182) RNA-binding glycine rich protein (RGP-2)

[Nicotiana sylvestris]

Seq. No.

6295

Contig ID 6478_4.R1040

5'-most EST LIB3094-002-Q1-K1-G1

Method BLASTX
NCBI GI g1084415
BLAST score 307
E value 4.0e-28
Match length 107
% identity 59

NCBI Description RNA-binding protein - Wood tobacco >gi 624925 dbj BAA05170_

(D26182) RNA-binding glycine rich protein (RGP-2)

[Nicotiana sylvestris]

Seq. No.

6296

Contig ID 5'-most EST

6488_1.R1040 kmv700740304.h1

Method BLASTX
NCBI GI g2500426
BLAST score 343
E value 9.0e-32
Match length 124
% identity 56

NCBI Description 30S RIBOSOMAL PROTEIN S9 >gi 1652399 dbj BAA17321 (D90905)



30S ribosomal protein S9 [Synechocystis sp.]

```
Seq. No.
                   6297
Contig ID
                   6489 1.R1040
5'-most EST
                   rca700997926.hl
Method
                   BLASTX
NCBI GI
                   q1931652
                   209
BLAST score
                   2.0e-16
E value
Match length
                   101
                   50
% identity
NCBI Description
                  (U95973) phosphatidylinositol-4-phosphate 5-kinase isolog
                   [Arabidopsis thaliana]
Seq. No.
                   6298
Contig ID
                   6498 1.R1040
5'-most EST
                  LIB3029-001-Q1-B1-H10
Seq. No.
                   6299
Contig ID
                   6500 1.R1040
5'-most EST
                   uC-gmronoir008d08b1
                  BLASTX
Method
NCBI GI
                   q3738320
BLAST score
                   459
                   3.0e-45
E value
                   303
Match length
% identity
NCBI Description
                   (AC005170) putative cinnamoyl CoA reductase [Arabidopsis
                   thaliana]
Seq. No.
                   6300
Contig ID
                   6500 2.R1040
                  LIB3051-010-Q1-E1-D4
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3738320
BLAST score
                   280
                   7.0e-25
E value
Match length
                   129
                   40
% identity
                   (AC005170) putative cinnamoyl CoA reductase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   6301
Contig ID
                   6520 1.R1040
5'-most EST
                  hyd700728002.h1
Method
                   BLASTX
NCBI GI
                   g2880051
BLAST score
                   227
E value
                   4.0e-18
Match length
                  101
% identity
                   47
NCBI Description
                  (AC002340) putative protein kinase [Arabidopsis thaliana]
```

Seq. No. 6302

Contig ID 6522 1.R1040 5'-most EST dpv701097360.h1

Method BLASTX



```
q4165018
NCBI GI
                   1037
BLAST score
                   1.0e-156
E value
                   615
Match length
% identity
                  (D89053) Acyl-CoA synthetase 3 [Homo sapiens]
NCBI Description
                   6303
Seq. No.
                   6522 2.R1040
Contig ID
                   LIB3029-001-Q1-B1-D4
5'-most EST
Method
                   BLASTX
                   g1468969
NCBI GI
                   172
BLAST score
                   6.0e-12
E value
                   56
Match length
% identity
NCBI Description (D30666) brain acyl-CoA synthtase II [Rattus norvegicus]
                   6304
Seq. No.
                   6522 3.R1040
Contig ID
                   uaw700662748.h1
5'-most EST
                   BLASTX
Method
                   q1468969
NCBI GI
                   380
BLAST score
                   1.0e-36
E value
                   134
Match length
                   57
% identity
                   (D30666) brain acyl-CoA synthtase II [Rattus norvegicus]
NCBI Description
                   6305
Seq. No.
                   6526 1.R1040
Contig ID
                   k11701206060.hl
5'-most EST
Method
                   BLASTX
                   q3482913
NCBI GI
BLAST score
                   646
E value
                   2.0e-67
                   253
Match length
% identity
                   (AC003970) Similar to MtN21, gi_2598575, Megicago
NCBI Description
                   truncatula nodulation induced gene [Arabidopsis thaliana]
                   6306
 Seq. No.
                   6532 1.R1040
 Contia ID
                   LIB3040-015-Q1-E1-C6
 5'-most EST
 Seq. No.
                   6307
                   6532 2.R1040
 Contig ID
                   LIB3040-053-Q1-E1-F12
 5'-most EST
                   6308
 Seq. No.
                   6533 1.R1040
 Contig ID
                   crh7\overline{0}0851849.h1
 5'-most EST
                   BLASTX
 Method
                   g265975
 NCBI GI
                   173
 BLAST score
```

8.0e-12

134

E value

Match length



% identity

NCBI Description RD29B=responsive-to-dessication protein [Arabidopsis

thaliana, Columbia ecotype, Peptide, 604 aa]

>gi 445066 prf 1908381B rd29B gene [Arabidopsis thaliana]

Seq. No. 6309

Contig ID 6533 2.R1040 5'-most EST awf700837810.hl

BLASTX Method NCBI GI q2654206 BLAST score 183 5.0e-13 E value Match length 295 19 % identity

NCBI Description (AF035535) cold acclimation protein [Spinacia oleracea] >gi 2661094 (AF035536) cold acclimation protein [Spinacia

oleracea]

Seq. No.

6310

Contig ID 6535 1.R1040

5'-most EST LIB3029-001-Q1-B1-E9

Seq. No.

Contig ID 6536 1.R1040

5'-most EST LIB3139-053-P1-N1-F11

6311

Seq. No. 6312

Contig ID 6538 1.R1040

5'-most EST uC-gmrominsoy172b01b1

Method BLASTX NCBI GI q2244852 BLAST score 615 E value 9.0e-64 195 Match length % identity

NCBI Description (Z97337) hypothetical protein [Arabidopsis thaliana]

Seq. No.

Contig ID 6538 3.R1040 5'-most EST zhf700957864.h1

6314 Seq. No.

Contig ID 6543 1.R1040

5'-most EST uC-gmrominsoy272d12b1

Method BLASTX NCBI GI g1871185 BLAST score 163 5.0e-11 E value Match length 44 % identity 68

NCBI Description (U90439) seven in absentia isolog [Arabidopsis thaliana]

Seq. No.

Contig ID 6544 1.R1040

5'-most EST LIB3029-001-Q1-B1-D1

6315

Method BLASTX NCBI GI g4539423



BLAST score 341 E value 4.0e-32 Match length 120 % identity 58

NCBI Description (AL049171) pyrophosphate-dependent phosphofructo-1-kinase

[Arabidopsis thaliana]

Seq. No. 6316

Contig ID 6545_1.R1040 5'-most EST gsv701051793.h1

Method BLASTX
NCBI GI g132659
BLAST score 732
E value 2.0e-77
Match length 241
% identity 60

NCBI Description 50S RIBOSOMAL PROTEIN L13, CHLOROPLAST PRECURSOR (CL13)

>gi_81483_pir__A32033 ribosomal protein L13 precursor, chloroplast - spinach >gi 170133 (J04461) ribosomal protein

L13 [Spinacia oleracea]

Seq. No. 6317

Contig ID 6551 1.R1040

5'-most EST uC-gmronoir007f09b1

Method BLASTX
NCBI GI g2911058
BLAST score 275
E value 2.0e-38
Match length 112
% identity 71

NCBI Description (AL021961) putative protein [Arabidopsis thaliana]

Seq. No. 6318

Contig ID 6551 2.R1040

5'-most EST jC-gmst02400002c05a1

Seq. No. 6319

Contig ID 6560 1.R1040

5'-most EST fC-gmse7000763420f1

Method BLASTX
NCBI GI g3413699
BLAST score 473
E value 3.0e-47
Match length 115
% identity 73

NCBI Description (AC004747) putative heat shock protein [Arabidopsis

thaliana]

Seq. No. 6320

Contig ID 6596 1.R1040

5'-most EST LIB3029-009-Q1-B1-F4

Seq. No. 6321

Contig ID 6598 1.R1040

5'-most EST LIB3170-086-Q1-J1-C3

Method BLASTN NCBI GI g2924257



```
BLAST score
                  1.0e-135
E value
                  880
Match length
                  76
% identity
NCBI Description Tobacco chloroplast genome DNA
Seq. No.
                  6322
                  6602 1.R1040
Contig ID
                  bth700847367.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q466530
BLAST score
                  266
                  1.0e-147
E value
                  878
Match length
                  83
% identity
NCBI Description Citrullus vulgaris mRNA for cysteine synthase, complete cds
                   6323
Seq. No.
                   6606 1.R1040
Contig ID
5'-most EST
                  LIB3139-118-P1-N1-F10
                  BLASTN
Method
NCBI GI
                  q4097879
                  251
BLAST score
                  1.0e-138
E value
                  815
Match length
                   86
% identity
                  Bean pod mottle virus complete segment RNA1 polyprotein
NCBI Description
                   gene, complete cds
                   6324
Seq. No.
                   6608 1.R1040
Contig ID
                   LIB3094-071-Q1-K1-E1
5'-most EST
Method
                   BLASTX
                   q1762945
NCBI GI
BLAST score
                   415
E value
                   3.0e-40
Match length
                   155
% identity
                   54
                  (U66269) ORF; able to induce HR-like lesions [Nicotiana
NCBI Description
                   tabacum]
                   6325
Seq. No.
Contig ID
                   6608 2.R1040
                   LIB3170-014-Q1-K1-G6
5'-most EST
                   BLASTX
Method
                   g1762945
NCBI GI
BLAST score
                   258
                   2.0e-44
E value
                   136
Match length
                   69
% identity
                  (U66269) ORF; able to induce HR-like lesions [Nicotiana
NCBI Description
```

tabacum]

Seq. No.

6326

Contig ID

6609 1.R1040

5'-most EST

LIB3109-002-Q1-K1-A10

Method

BLASTX



```
g4056420
NCBI GI
BLAST score
                  454
                  8.0e-45
E value
                  164
Match length
                  53
% identity
                  (AC005322) ESTs gb T144077 and gb T43352 come from this
NCBI Description
                  gene. [Arabidopsis thaliana]
                  6327
Seq. No.
                  6609 2.R1040
Contig ID
5'-most EST
                  LIB3039-047-Q1-E1-B9
Method
                  BLASTX
                  g3298546
NCBI GI
                  175
BLAST score
                  1.0e-12
E value
Match length
                  61
% identity
                  (AC004681) unknown protein [Arabidopsis thaliana]
NCBI Description
                  6328
Seq. No.
                  6612 1.R1040
Contig ID
5'-most EST
                  LIB3109-037-Q1-K1-G5
                  BLASTX
Method
                  g2191149
NCBI GI
BLAST score
                  351
                  2.0e-39
E value
Match length
                  188
                  43
% identity
                  (AF007269) Similar to protein kinase [Arabidopsis thaliana]
NCBI Description
                  6329
Seq. No.
Contig ID
                  6625 1.R1040
5'-most EST
                  LIB3029-011-Q1-B1-G12
Seq. No.
                  6330
                  6629 1.R1040
Contig ID
5'-most EST
                  uaw700664958.h1
                  BLASTX
Method
                  g3763916
NCBI GI
BLAST score
                  274
                   4.0e-24
E value
Match length
                  164
% identity
                   37
                  (AC004450) unknown protein [Arabidopsis thaliana]
NCBI Description
                  >gi 4531439 gb AAD22124.1 AC006224 6 (AC006224) unknown
                  protein [Arabidopsis thaliana]
                   6331
Seq. No.
                   6632 1.R1040
Contig ID
5'-most EST
                   g5510323
                   BLASTX
Method
NCBI GI
                   g4049518
BLAST score
                   463
E value
                   9.0e - 46
```

E value 9.0e-46
Match length 230
% identity 42
NCBI Description (ALO31852) conserved

NCBI Description (AL031852) conserved hypothetical protein



[Schizosaccharomyces pombe]

```
Seq. No.
                  6332
Contig ID
                  6637 1.R1040
5'-most EST
                  hyd700725528.h1
Seq. No.
                  6333
Contia ID
                  6638 1.R1040
5'-most EST
                  jC-gmle01810021b02a1
Method
                  BLASTX
NCBI GI
                  g3080407
BLAST score
                  424
E value
                  1.0e-84
Match length
                  221
% identity
NCBI Description
                  (AL022604) hypothetical protein [Arabidopsis thaliana]
Seq. No.
Contig ID
                  6638 2.R1040
5'-most EST
                  uC-qmflminsoy089f09b1
Method
                  BLASTX
NCBI GI
                  q3080407
BLAST score
                  435
E value
                  5.0e-43
Match length
                  138
                  79
% identity
NCBI Description
                  (AL022604) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  6335
Contig ID
                  6639 1.R1040
5'-most EST
                  gsv701049521.hl
Method
                  BLASTX
NCBI GI
                  g461736
BLAST score
                  1763
E value
                  0.0e+00
Match length
                  385
% identity
                  91
                  MITOCHONDRIAL CHAPERONIN HSP60-2 PRECURSOR
NCBI Description
                  >gi 478786 pir S29316 chaperonin 60 - cucurbit
                  >gi 12546 emb CAA50218 (X70868) chaperonin 60 [Cucurbita
                  sp.]
                  6336
Seq. No.
                  6643 1.R1040
Contig ID
5'-most EST
                  LIB3107-013-Q1-K1-E2
Method
                  BLASTX
NCBI GI
                  g3738327
BLAST score
                  374
E value
                  1.0e-35
Match length
                  148
% identity
                  49
NCBI Description
                  (AC005170) putative serine carboxypeptidase [Arabidopsis
                  thaliana]
```

Seq. No. 6337

Contig ID 6643_2.R1040

5'-most EST jC-gmst02400029a10a1



```
Method
                   BLASTX
NCBI GI
                   q3738327
BLAST score
                   365
                   7.0e-35
E value
Match length
                   119
% identity
                   57
                   (AC005170) putative serine carboxypeptidase [Arabidopsis
NCBI Description
                   6338
Seq. No.
Contig ID
                   6644 1.R1040
                   hyd7\overline{0}0730480.h1
5'-most EST
                   BLASTX
Method
                   q3935152
NCBI GI
BLAST score
                   1752
E value
                   0.0e+00
Match length
                   413
% identity
                   (AC005106) T25N20.16 [Arabidopsis thaliana]
NCBI Description
                   6339
Seq. No.
Contig ID
                   6652 1.R1040
5'-most EST
                   LIB3170-074-Q1-K1-E10
                   6340
Seq. No.
Contig ID
                   6652 2.R1040
                   jC-gmle01810088e06a1
5'-most EST
Seq. No.
                   6341
                   6654 1.R1040
Contig ID
                   LIB3030-012-Q1-B1-D11
5'-most EST
Method
                   BLASTX
                   q3582343
NCBI GI
                   167
BLAST score
                   9.0e-12
E value
                   100
Match length
% identity
                   34
                   (AC005496) putative flavonol 3-o-glucosyltransferase
NCBI Description
                    [Arabidopsis thaliana]
                   6342
Seq. No.
                   6655 1.R1040
Contig ID
                   pcp7\overline{0}0994725.h1
5'-most EST
                   BLASTX
Method
                   g2961384
NCBI GI
BLAST score
                   618
                   5.0e-70
E value
                   254
Match length
                   52
% identity
```

(AL022141) aldehyde dehydrogenase like protein [Arabidopsis NCBI Description

thaliana]

6343 Seq. No. Contig ID 6656 1.R1040 $q439\overline{7}369$ 5'-most EST

Method BLASTX NCBI GI g4539007



```
BLAST score
                  1.0e-57
E value
Match length
                  112
                  91
% identity
NCBI Description (AL049481) cytochrome c [Arabidopsis thaliana]
                  6344
Seq. No.
Contig ID
                  6656 2.R1040
5'-most EST
                  LIB3039-029-Q1-E1-F11
                  BLASTX
Method
                  q118013
NCBI GI
BLAST score
                  579
                  7.0e-60
E value
                  111
Match length
% identity
NCBI Description CYTOCHROME C >gi 65497 pir CCMB cytochrome c - mung bean
Seq. No.
                  6345
Contig ID
                  6661 1.R1040
5'-most EST
                  LIB3093-003-Q1-K1-B6
Method
                  BLASTX
NCBI GI
                  q4204912
BLAST score
                  1215
E value
                  1.0e-134
Match length
                  396
                   61
% identity
NCBI Description (U58918) MEK kinase [Arabidopsis thaliana]
Seq. No.
                   6346
Contig ID
                   6661 2.R1040
5'-most EST
                  LIB3139-095-P1-N1-H3
Method
                  BLASTX
                   q4204912
NCBI GI
BLAST score
                   225
                   5.0e-18
E value
                  129
Match length
% identity
                   44
                  (U58918) MEK kinase [Arabidopsis thaliana]
NCBI Description
                   6347
Seq. No.
                   6661 3.R1040
Contig ID
5'-most EST
                   jC-gmle01810063e05d1
                   6348
Seq. No.
                   6661 7.R1040
Contig ID
5'-most EST
                  LIB3170-024-Q1-J1-A4
                   6349
Seq. No.
                   6670 1.R1040
Contig ID
5'-most EST
                   LIB3030-012-Q1-B1-A8
Method
                   BLASTX
                   q996020
NCBI GI
                   387
BLAST score
E value
                   2.0e-66
                   237
Match length
% identity
                   53
```

1190

NCBI Description (X91638) BRM protein [Gallus gallus]



```
6350
Seq. No.
                   6672 1.R1040
Contig ID
5'-most EST
                  LIB3030-012-Q1-B1-B10
                   6351
Seq. No.
Contig ID
                   6673 1.R1040
                  uC-gmflminsoy024d02b1
5'-most EST
Seq. No.
                   6352
                   6678_1.R1040
Contig ID
                  sat7\overline{0}1014952.h1
5'-most EST
                   6353
Seq. No.
Contig ID
                   6683 1.R1040
5'-most EST
                   rlr700900487.hl
Method
                   BLASTX
                   q2708532
NCBI GI
BLAST score
                   970
                   1.0e-105
E value
Match length
                   301
% identity
                   32
                  (AF029351) putative RNA binding protein [Nicotiana tabacum]
NCBI Description
                   6354
Seq. No.
Contig ID
                   6687 1.R1040
5'-most EST
                   jC-gmle01810000g09a1
Seq. No.
                   6355
                   6687 2.R1040
Contig ID
5'-most EST
                   zpv700760473.hl
                   6356
Seq. No.
Contig ID
                   6694 1.R1040
5'-most EST
                   LIB3030-011-Q1-B1-H10
Method
                   BLASTX
                   g1504020
NCBI GI
BLAST score
                   122
                   6.0e-10
E value
Match length
                   55
% identity
                   67
                  (D86973) similar to Yeast translation activator GCN1
NCBI Description
                   (P1:A48126) [Homo sapiens]
                   6357
Seq. No.
Contig ID
                   6699 1.R1040
```

5'-most EST LIB3051-103-Q1-K1-G7

Method BLASTX NCBI GI q461812 BLAST score 414 E value 2.0e-40 Match length 193 % identity

CYTOCHROME P450 72 (CYPLXXII) (PROBABLE NCBI Description

GERANIOL-10-HYDROXYLASE) (GE10H) >gi 167484 (L10081)

Cytochrome P-450 protein [Catharanthus roseus]

>gi 445604 prf 1909351A cytochrome P450 [Catharanthus



roseus]

152

49

Match length

NCBI Description

% identity

```
Seq. No.
                   6358
                   6699 2.R1040
Contig ID
5'-most EST
                   uC-gmronoir043h10b1
Seq. No.
                   6359
                   6699 5.R1040
Contig ID
                   zhf700952542.hl
5'-most EST
                   6360
Seq. No.
                   6699 6.R1040
Contig ID
                   zhf700961126.hl
5'-most EST
                   6361
Seq. No.
                   6703 1.R1040
Contig ID
5'-most EST
                   LIB3030-012-Q1-B1-A12
                   BLASTX
Method
NCBI GI
                   g4467124
                   172
BLAST score
                   3.0e-12
E value
                   107
Match length
                   43
% identity
                  (AL035538) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   6362
Seq. No.
                   6703 2.R1040
Contig ID
                   jC-gmf102220102h03a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4467124
                   159
BLAST score
                   1.0e-10
E value
Match length
                   43
                   70
% identity
                  (AL035538) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   6363
Seq. No.
                   6704 1.R1040
Contig ID
                   LIB3030-011-Q1-B1-B11
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4006915
BLAST score
                   215
E value
                   3.0e-17
Match length
                   111
% identity
                   (Z99708) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6364
                   6706 1.R1040
Contig ID
                   zhf7\overline{0}0965022.h1
5'-most EST
                   BLASTX
Method
                   q4415908
NCBI GI
BLAST score
                   431
                   3.0e-42
E value
```

(AC006282) unknown protein [Arabidopsis thaliana]



Seq. No. 6365

Contig ID 6713_1.R1040

5'-most EST uC-gmflminsoy119a09b1

Method BLASTX
NCBI GI g3080437
BLAST score 859
E value 3.0e-92
Match length 379
% identity 54

NCBI Description (AL022605) putative protein [Arabidopsis thaliana]

Seq. No. 6366

Contig ID 6715_1.R1040 5'-most EST xpa700797631.h1

Seq. No. 6367

Contig ID 6715_2.R1040

5'-most EST LIB3040-043-Q1-E1-G1

Seq. No. 6368

Contig ID 6716_1.R1040

5'-most EST LIB3051-079-Q1-K1-C3

Method BLASTN
NCBI GI 9488576
BLAST score 204
E value 1.0e-111
Match length 416
% identity 87

% identity 87
NCBI Description Medicago sativa Regen S clone pH3c131 histone H3.2 mRNA,

complete cds

Seq. No. 6369

Contig ID 6716_2.R1040

5'-most EST LIB3049-029-Q1-E1-D2

Method BLASTN
NCBI GI g488576
BLAST score 193
E value 1.0e-104
Match length 413
% identity 87

NCBI Description Medicago sativa Regen S clone pH3c131 histone H3.2 mRNA,

complete cds

Seq. No. 6370

Contig ID 6716_3.R1040

5'-most EST LIB3049-019-Q1-E1-H7

Method BLASTN
NCBI GI g1208641
BLAST score 81
E value 1.0e-37
Match length 197
% identity 93

NCBI Description Amphicarpa bracteata histone H3 gene, partial cds, clone

H3-B2AMC

Seq. No. 6371



```
6716 4.R1040
Contig ID
5'-most EST
                  LIB3170-062-Q1-J1-G11
                   BLASTN
Method
                   g488576
NCBI GI
BLAST score
                   65
E value
                   8.0e-28
                   139
Match length
                   91
% identity
                  Medicago sativa Regen S clone pH3c131 histone H3.2 mRNA,
NCBI Description
                   complete cds
                   6372
Seq. No.
                   6720 1.R1040
Contig ID
                   LIB3030-011-Q1-B1-E12
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2618691
BLAST score
                   449
                   6.0e-60
E value
Match length
                   185
                   64
% identity
                   (AC002510) putative chloroplast envelope Ca2+-ATPase
NCBI Description
                   [Arabidopsis thaliana]
                   6373
Seq. No.
                   6722 1.R1040
Contig ID
                   pmv700895035.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3643594
BLAST score
                   557
                   1.0e-56
E value
                   249
Match length
% identity
                   (AC005395) unknown protein [Arabidopsis thaliana]
NCBI Description
                   6374
Seq. No.
                   6722 2.R1040
Contig ID
5'-most EST
                   LIB3050-023-Q1-K1-F9
Method
                   BLASTX
NCBI GI
                   q3643594
BLAST score
                   539
E value
                   2.0e-98
Match length
                   246
% identity
                   (AC005395) unknown protein [Arabidopsis thaliana]
NCBI Description
                   6375
Seq. No.
                   6722 3.R1040
Contig ID
                   LIB3030-010-Q1-B1-G6
5'-most EST
                   BLASTX
Method
                   g3643594
NCBI GI
                   271
BLAST score
                   1.0e-23
E value
Match length
                   63
                   83
% identity
```

NCBI Description (AC

6376

Seq. No.

(AC005395) unknown protein [Arabidopsis thaliana]



Contig ID 6726_1.R1040 5'-most EST pmv700892405.h1

Method BLASTX
NCBI GI g4455369
BLAST score 323
E value 1.0e-29
Match length 174
% identity 41

NCBI Description (AL035524) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6377

Contig ID 6732 1.R1040

5'-most EST uC-gmflminsoy115g08b1

Seq. No. 6378

Contig ID 6734 1.R1040

5'-most EST jC-gmfl02220106h03a1

Method BLASTX
NCBI GI g2911059
BLAST score 909
E value 3.0e-98
Match length 197
% identity 84

NCBI Description (AL021961) putative protein [Arabidopsis thaliana]

Seq. No. 6379

Contig ID 6736_1.R1040 5'-most EST pxt700943944.h1

Method BLASTX
NCBI GI g1076668
BLAST score 1095
E value 1.0e-129
Match length 243
% identity 93

NCBI Description NADH dehydrogenase (EC 1.6.99.3) - potato

>gi_639834_emb_CAA58823_ (X83999) NADH dehydrogenase

[Solanum tuberosum]

Seq. No. 6380

Contig ID 6737 1.R1040

5'-most EST LIB3030-010-Q1-B1-E9

Seq. No. 6381

Contig ID 6738_1.R1040

5'-most EST uC-gmropic036c09b1

Method BLASTN
NCBI GI g2337888
BLAST score 35
E value 6.0e-10
Match length 139
% identity 81

NCBI Description Genomic sequence for Arabidopsis thaliana BAC F14J16,

complete sequence [Arabidopsis thaliana]

Seq. No. 6382

Contig ID 6741 1.R1040 5'-most EST kl1701213610.h1



```
BLASTX
Method
                  q3876874
NCBI GI
                  344
BLAST score
                  9.0e - 32
E value
                  408
Match length
                  28
% identity
NCBI Description (Z81535) Similarity to Human signal recognition particle
                  receptor alpha subunit (SW:P08240) [Caenorhabditis elegans]
Seq. No.
                  6742 1.R1040
Contig ID
                  LIB3030-010-Q1-B1-F7
5'-most EST
                  BLASTX
Method
                  g3395584
NCBI GI
                  291
BLAST score
                  5.0e-34
E value
                  147
Match length
                   54
% identity
                  (AL031179) importin beta subunit [Schizosaccharomyces
NCBI Description
                  pombe]
                   6384
Seq. No.
                   6744 1.R1040
Contig ID
                   LIB3030-010-Q1-B1-G1
5'-most EST
                   BLASTX
Method
                   g2055273
NCBI GI
                   1716
BLAST score
                   0.0e+00
E value
                   386
Match length
                   84
% identity
                  (D85339) hydroxypyruvate reductase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6385
                   6745 1.R1040
Contig ID
                   q4290785
5'-most EST
                   6386
Seq. No.
                   6745 2.R1040
Contig ID
                   uC-gmflminsoy002h05b1
5'-most EST
                   6387
Seq. No.
Contig ID
                   6746 1.R1040
                   leu701156386.hl
 5'-most EST
Method
                   BLASTX
                   q4056467
NCBI GI
BLAST score
                   335
                   4.0e-31
E value
                   126
Match length
 % identity
                   (AC005990) Strong similarity to gb_AB006693 spermidine
 NCBI Description
                   synthase from Arabidopsis thaliana. ESTs gb_AA389822,
                   gb_T41794, gb_N38455, gb_AI100106, gb_F14442 and gb_F14256
                   come from this gene. [Arabidopsis thaliana]
```

Seq. No. 6388

Contig ID 6746_3.R1040

5'-most EST LIB3040-035-Q1-E1-B3



```
6389
Seq. No.
                   6753 1.R1040
Contig ID
                  kl1701206385.h1
5'-most EST
                  BLASTX
Method
                   g3819099
NCBI GI
                   152
BLAST score
E value
                   1.0e-09
Match length
                  84
% identity
                   42
NCBI Description (AJ009825) copper amine oxidase [Cicer arietinum]
                   6390
Seq. No.
                   6756 1.R1040
Contig ID
                   LIB3T70-003-Q1-K1-B4
5'-most EST
                   BLASTX
Method
                   g2769642
NCBI GI
BLAST score
                   760
                   1.0e-80
E value
                   224
Match length
                   64
% identity
                   (Z97215) nine-cis-epoxycarotenoid dioxygenase [Lycopersicon
NCBI Description
                   esculentum]
                   6391
Seq. No.
                   6757 1.R1040
Contig ID
                   LIB3\overline{1}07-079-Q1-K1-B11
5'-most EST
                   6392
Seq. No.
                   6757 2.R1040
Contig ID
5'-most EST
                   jex700906928.h1
                   6393
Seq. No.
                   6758 1.R1040
Contig ID
                   jC-gmle01810085e04a1
5'-most EST
                   6394
Seq. No.
                   6762 1.R1040
Contig ID
                   jC-gmst02400014g11a1
5'-most EST
                   BLASTX
Method
                   q3337366
NCBI GI
BLAST score
                   413
E value
                   3.0e-40
Match length
                   199
% identity
                   (AC004481) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6395
Contig ID
                    6763 1.R1040
                   ncj700982144.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   g231610
```

Method BLASIX NCBI GI g231610 BLAST score 1445 E value 1.0e-161 Match length 374 % identity 78

NCBI Description ATP SYNTHASE GAMMA CHAIN, CHLOROPLAST PRECURSOR



>gi_67880_pir__PWNTG H+-transporting ATP synthase (EC
3.6.1.34) gamma chain precursor, chloroplast - common
tobacco >gi_19785_emb_CAA45152_ (X63606) ATP synthase
(gamma subunit) [Nicotiana tabacum]

6396 Seq. No. 6765 1.R1040 Contig ID jsh701066574.h15'-most EST BLASTX Method q4337175 NCBI GI 510 BLAST score 2.0e-51 E value 137 Match length 69 % identity (AC006416) ESTs gb_T20589, gb_T04648, gb_AA597906, NCBI Description gb_T04111, gb_R84180, gb_R65428, gb_T44439, gb_T76570, gb_R90004, gb_T45020, gb_T42457, gb_T20921, gb_AA042762 and gb_AA720210 come from this gene. [Arabidopsis thaliana] Seq. No. 6766 1.R1040 Contig ID 5'-most EST q5058357 Method BLASTN g3873174 NCBI GI 35 BLAST score 9.0e-10 E value 179 Match length % identity Genomic sequence for Arabidopsis thaliana BAC F14N23, NCBI Description complete sequence [Arabidopsis thaliana] 6398 Seq. No. 6766 2.R1040 Contig ID LIB3093-011-Q1-K1-D4 5'-most EST 6399 Seq. No. 6766 3.R1040 Contig ID jC-gmro02910047f09a1 5'-most EST BLASTX Method q4191616 NCBI GI BLAST score 588 7.0e-61 E value Match length 157 73 % identity (AF120334) GTP-binding protein NGB [Homo sapiens] NCBI Description

Seq. No. 6400

Contig ID 6766_4.R1040

5'-most EST uC-gmrominsoy229a04b1

Method BLASTX
NCBI GI g4191616
BLAST score 377
E value 2.0e-36
Match length 93
% identity 69

NCBI Description (AF120334) GTP-binding protein NGB [Homo sapiens]

Method

NCBI GI

BLAST score



```
6401
Seq. No.
                  6766 5.R1040
Contig ID
                  leu701149247.h1
5'-most EST
                  BLASTX
Method
                  q4191616
NCBI GI
BLAST score
                  268
                  2.0e-23
E value
                  70
Match length
% identity
NCBI Description (AF120334) GTP-binding protein NGB [Homo sapiens]
Seq. No.
                   6766 8.R1040
Contig ID
                   smc700744313.h1
5'-most EST
                   6403
Seq. No.
Contig ID
                   6766 10.R1040
5'-most EST
                   gsv701044187.h1
                   BLASTX
Method
                   q4191616
NCBI GI
                   154
BLAST score
E value
                   2.0e-10
Match length
                   42
                   67
% identity
NCBI Description (AF120334) GTP-binding protein NGB [Homo sapiens]
Seq. No.
                   6404
                   6768 1.R1040
Contig ID
                   LIB3092-044-Q1-K1-D7
5'-most EST
Method
                   BLASTX
                   q3023271
NCBI GI
BLAST score
                   1511
E value
                   0.0e + 00
                   381
Match length
% identity
                   GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE (FDH)
NCBI Description
                   (FALDH) (GSH-FDH) >gi_1675394 (U77637) class III ADH enzyme
                   [Oryza sativa]
                   6405
Seq. No.
Contig ID
                   6768 2.R1040
                   LIB3109-028-Q1-K1-A7
5'-most EST
Method
                   BLASTX
                   g1841502
NCBI GI
                   453
BLAST score
                   2.0e-47
E value
                   105
Match length
% identity
                   90
                   (Y11029) glutothione-dependent formaldehyde dehydrogenase
NCBI Description
                   [Zea mays]
                   6406
Seq. No.
                   6769 1.R1040
Contig ID
                   LIB3030-010-Q1-B1-C8
5'-most EST
                   BLASTX
```

g3925363

344



```
E value
                   2.0e-32
Match length
                   96
% identity
                   79
NCBI Description
                   (AF067961) homeodomain protein [Malus domestica]
Seq. No.
                   6407
                   6774 1.R1040
Contig ID
5'-most EST
                  gsv701056309.h1
Seq. No.
                   6408
                   6774 2.R1040
Contig ID
5'-most EST
                  hrw701060177.h1
Seq. No.
                   6409
Contig ID
                   6775 1.R1040
5'-most EST
                   fde700876739.hl
Method
                  BLASTX
NCBI GI
                  g4522012
BLAST score
                  899
E value
                   7.0e-97
                  310
Match length
% identity
                   57
NCBI Description
                  (AC007069) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   6410
Contig ID
                   6778 1.R1040
5'-most EST
                  pxt700946334.h1
Method
                  BLASTX
NCBI GI
                   q3063470
BLAST score
                   396
                   2.0e-38
E value
Match length
                  70
% identity
                   97
NCBI Description
                  (AC003981) F22013.32 [Arabidopsis thaliana]
Seq. No.
                   6411
Contig ID
                   6782 1.R1040
5'-most EST
                   uC-qmrominsoy243g03b1
Seq. No.
                   6412
                   6782 2.R1040
Contig ID
5'-most EST
                   awf700840314.hl
Seq. No.
                   6413
Contig ID
                   6782 3.R1040
5'-most EST
                  uC-gmflminsoy078f04b1
                   6414
Seq. No.
                   6783 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy142c04b1
                   BLASTX
Method
```

NCBI GI g3540219 BLAST score 891 E value 5.0e-96 Match length 390 % identity 46

NCBI Description (D87686) KIAA0017 protein [Homo sapiens]



```
Seq. No.
                  6415
Contig ID
                  6784 1.R1040
5'-most EST
                  uaw700661165.hl
Method
                  BLASTX
NCBI GI
                  g322787
BLAST score
                  755
                  7.0e-80
E value
                  348
Match length
                  46
% identity
NCBI Description pyruvate kinase (EC 2.7.1.40), cytosolic - potato
Seq. No.
                  6416
                  6784 2.R1040
Contig ID
5'-most EST
                  jC-qmst02400072a04a1
Seq. No.
                  6417
                  6785 1.R1040
Contig ID
5'-most EST
                  LIB3030-010-Q1-B1-B3
Seq. No.
                  6418
                  6787 1.R1040
Contig ID
5'-most EST
                  LIB3030-009-Q1-B1-G5
Method
                  BLASTX
NCBI GI
                  g3292830
BLAST score
                  346
E value
                  4.0e-32
Match length
                  263
% identity
NCBI Description (AL031018) putative protein [Arabidopsis thaliana]
Seq. No.
                  6419
Contig ID
                  6791 1.R1040
5'-most EST
                  uC-gmflminsoy018e10b1
Method
                  BLASTX
NCBI GI
                  q2828151
BLAST score
                  302
E value
                  2.0e-27
Match length
                  100
                  59
% identity
                  (AF042386) cyclophilin-33B [Homo sapiens]
NCBI Description
                  >gi_4406229_gb_AAD19907_ (AF104013) peptidyl-prolyl
                  cis-trans isomerase E [Homo sapiens]
                  6420
Seq. No.
Contig ID
                  6796 1.R1040
5'-most EST
                  LIB3030-009-Q1-B1-H7
Method
                  BLASTX
NCBI GI
                  g4263704
BLAST score
                  526
E value
                  3.0e-53
Match length
                  239
% identity
                  43
NCBI Description
                  (AC006223) putative sugar starvation-induced protein
```

Seq. No. 6421

[Arabidopsis thaliana]



Contig ID 6796 3.R1040 5'-most EST zhf700958381.h1 Method BLASTX NCBI GI q4263704 BLAST score 163 E value 4.0e-11 Match length 64 42 % identity (AC006223) putative sugar starvation-induced protein NCBI Description [Arabidopsis thaliana] 6422 Seq. No. 6797 1.R1040 Contig ID LIB3109-048-Q1-K1-B1 5'-most EST BLASTX Method NCBI GI g4105782 BLAST score 986 E value 1.0e-107 215 Match length % identity 85 (AF049922) PGP169-12 [Petunia x hybrida] NCBI Description Seq. No. 6423 6799 1.R1040 Contig ID 5'-most EST $q429\overline{1}450$ Method BLASTN NCBI GI g3241924 BLAST score 37 E value 1.0e-10 Match length 128 % identity NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MNC6, complete sequence [Arabidopsis thaliana] Seq. No. 6424 6800 1.R1040 Contig ID 5'-most EST LIB3030-009-Q1-B1-F2 Method BLASTX q626009 NCBI GI BLAST score 800 E value 1.0e-85 200 Match length % identity 71 capsanthin-capsorubin synthase (EC 1.-.-.) - pepper NCBI Description >gi 468748 emb CAA54495 (X77289) capsanthin/capsorubin synthase [Capsicum annuum] >gi 522120_emb_CAA53759_ (X76165) capsanthin/capsorubin sythase [Capsicum annuum] 6425 Seq. No. Contig ID 6804 2.R1040 5'-most EST wrg700789208.h2 Method BLASTX

Method BLASTX
NCBI GI g1922242
BLAST score 235
E value 1.0e-19
Match length 60
% identity 85



```
(Y10084) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   6426
Seq. No.
                   6804 3.R1040
Contig ID
5'-most EST
                   zpv700763346.h1
Method
                   BLASTX
                   q1922242
NCBI GI
BLAST score
                   210
                   1.0e-16
E value
                   53
Match length
% identity
                  (Y10084) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6427
                   6805 1.R1040
Contig ID
5'-most EST
                   LIB3030-009-Q1-B1-G1
Method
                   BLASTX
NCBI GI
                   q2462828
                   298
BLAST score
                   5.0e-27
E value
                   118
Match length
% identity
                   46
                  (AF000657) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   6428
Seq. No.
                   6805 2.R1040
Contig ID
5'-most EST
                   zsq701124373.h1
                   BLASTX
Method
                   g2462828
NCBI GI
BLAST score
                   167
                   6.0e-12
E value
Match length
                   59
% identity
                   47
                  (AF000657) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   6429
Seq. No.
                   6806 1.R1040
Contig ID
                   jC-gmro02910014f01a1
5'-most EST
                   BLASTX
Method
                   g4508083
NCBI GI
BLAST score
                   500
                   3.0e-50
E value
                   307
Match length
% identity
                   (AC005882) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6430
                   6807 1.R1040
Contig ID
5'-most EST
                   LIB3170-007-Q1-K1-D10
                   BLASTX
Method
NCBI GI
                   g3687251
BLAST score
                   542
                   3.0e-55
E value
Match length
                   142
                   73
 % identity
```

NCBI Description (AC005169) unknown protein [Arabidopsis thaliana]



```
Seq. No.
                  6810 1.R1040
Contig ID
                  LIB3049-001-Q1-E1-G8
5'-most EST
                  BLASTX
Method
                  g2213625
NCBI GI
BLAST score
                  478
E value
                  1.0e-47
                  169
Match length
                  54
% identity
NCBI Description (AC000103) F21J9.17 [Arabidopsis thaliana]
                  6432
Seq. No.
                  6810 2.R1040
Contig ID
                  LIB3030-009-Q1-B1-D9
5'-most EST
                  6433
Seq. No.
                  6810 4.R1040
Contig ID
                  pxt700941967.h1
5'-most EST
                   6434
Seq. No.
                   6811 1.R1040
Contig ID
                  uC-qmflminsoy008h02b1
5'-most EST
                  BLASTX
Method
                  g232024
NCBI GI
                  234
BLAST score
                   6.0e-19
E value
                   253
Match length
% identity
NCBI Description PROTEIN E6 >gi_421806_pir__A46130 fiber protein - upland
                   cotton >gi_2129498_pir__S65061 fiber protein E6 (clone
                   CKE6-1A) - upland cotton >gi_167323 (M92051) 5' start site
                   is putative; putative [Gossypium hirsutum] >gi_1000084
                   (U30505) E6 [Gossypium hirsutum]
                   6435
Seq. No.
                   6811 2.R1040
Contig ID
5'-most EST
                   g4295094
Method
                   BLASTX
                   q167324
NCBI GI
                   170
BLAST score
                   1.0e-11
E value
                   191
Match length
% identity
                  (M92051) 5' start site is putative; putative [Gossypium
NCBI Description
                   hirsutum]
                   6436
Seq. No.
                   6812 1.R1040
 Contig ID
                   LIB3065-008-Q1-N1-F6
 5'-most EST
                   BLASTX
Method
                   g2245080
 NCBI GI
 BLAST score
                   151
```

NCBI Description (Z97343) myosin heavy chain homolog [Arabidopsis thaliana]

9.0e-10

80

45

E value

Match length

% identity



```
6437
Seq. No.
                   6815 1.R1040
Contig ID
                   leu701152947.h1
5'-most EST
                   BLASTX
Method
                   g3273202
NCBI GI
                   345
BLAST score
                   3.0e - 32
E value
                   121
Match length
                   60
% identity
                  (AB010918) responce reactor4 [Arabidopsis thaliana]
NCBI Description
                   6438
Seq. No.
                   6819 1.R1040
Contig ID
                   leu701150644.hl
5'-most EST
                   6439
Seq. No.
                   6819 2.R1040
Contig ID
                   trc700561862.h1
5'-most EST
                   6440
Seq. No.
                   6819 3.R1040
Contig ID
                   LIB3055-007-Q1-N1-F11
5'-most EST
                   6441
Seq. No.
                   6820 1.R1040
Contig ID
                   uC-g\overline{m}rominsoy034b03b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3935157
BLAST score
                   1255
                   1.0e-139
E value
                   274
Match length
% identity
                  (AC005106) T25N20.21 [Arabidopsis thaliana]
NCBI Description
                   6442
Seq. No.
                   6821 1.R1040
Contig ID
                   uC-gmrominsoy201e01b1
5'-most EST
Method
                   BLASTX
                   g2982293
NCBI GI
                   152
BLAST score
E value
                   1.0e-19
                   81
Match length
% identity
                   59
                   (AF051231) ISP42-like protein [Picea mariana]
NCBI Description
                    6443
Seq. No.
                    6822 1.R1040
Contig ID
                   ncj700988638.hl
5'-most EST
                   BLASTX
Method
                    g2098705
NCBI GI
                    494
BLAST score
                    9.0e-50
E value
                    137
Match length
                    70
 % identity
                   (U82973) pectinesterase [Citrus sinensis]
```

Seq. No. 6444

NCBI Description



6830 1.R1040 Contig ID pxt700945726.h1 5'-most EST BLASTN Method g1223921 NCBI GI 127 BLAST score 5.0e-65 E value 331 Match length % identity 85 Vigna radiata vicilin peptidohydrolase, cysteinyl NCBI Description endopeptidase mRNA, complete cds 6445 Seq. No. 6832 1.R1040 Contig ID 5'-most EST jC-qmle01810024f07a1 6446 Seq. No. 6832 2.R1040 Contig ID LIB3030-009-Q1-B1-B7 5'-most EST 6447 Seq. No. 6832 3.R1040 Contig ID LIB3138-120-Q1-N1-E4 5'-most EST 6448 Seq. No. 6835 1.R1040 Contig ID LIB3030-009-Q1-B1-A6 5'-most EST BLASTX Method NCBI GI g632149 BLAST score 366 2.0e-34 E value 256 Match length % identity narbonin - Vicia pannonica >gi_396829_emb_CAA80981_ NCBI Description (Z25534) narbonin [Vicia pannonica] 6449 Seq. No. 6836 1.R1040 Contig ID 5'-most EST zhf700965238.h1 BLASTX Method q1362162 NCBI GI BLAST score 184 E value 7.0e-39 Match length 173 % identity 53 beta-glucosidase BGQ60 precursor - barley >gi_804656 NCBI Description (L41869) beta-glucosidase [Hordeum vulgare] 6450 Seq. No. 6841 2.R1040 Contig ID LIB3139-068-P1-N1-H10 5'-most EST 6451 Seq. No.

Contig ID 6844_1.R1040

5'-most EST LIB3170-053-Q1-K1-A5

Method BLASTX
NCBI GI g1330401
BLAST score 206



5.0e-16 E value Match length 118 % identity 36

NCBI Description (U58762) T27F7.1 gene product [Caenorhabditis elegans]

Seq. No.

Contig ID

6844 3.R1040

5'-most EST

LIB3030-008-Q1-B1-H9

Seq. No.

Contig ID

6845 1.R1040

5'-most EST

LIB3092-010-Q1-K1-E11

Seq. No.

6454

Contig ID

6846_1.R1040

5'-most EST

LIB3\overline{1}07-014-Q1-K1-G8

Method BLASTX NCBI GI g2132930 BLAST score 395 E value 4.0e-38 Match length 155 49

% identity

NCBI Description probable membrane protein YOR262w - yeast (Saccharomyces

cerevisiae) >gi 1420591 emb CAA99484 (Z75170) ORF YOR262w

[Saccharomyces cerevisiae]

Seq. No. 6455

Contig ID 6847 1.R1040

5'-most EST LIB3030-008-Q1-B1-H10

Method BLASTX NCBI GI g2244792 BLAST score 172 E value 8.0e-12 Match length 151 % identity 34

NCBI Description (Z97336) ankyrin homolog [Arabidopsis thaliana]

Seq. No. 6456

Contig ID 6848 1.R1040

5'-most EST LIB3030-008-Q1-B1-F4

BLASTX Method NCBI GI q1483218 BLAST score 369 E value 3.0e-35 Match length 164 % identity 50

NCBI Description (X99793) induced upon wounding stress [Arabidopsis

thaliana]

6457 Seq. No.

6850 1.R1040 Contig ID 5'-most EST kll701203944.h2

Method BLASTX NCBI GI q3193301 BLAST score 677 E value 6.0e-71 Match length 247



% identity (AF069298) Arabidopsis putative chloroplast outer envelope NCBI Description 86-like protein T10P11.19 (GB: AC002330) [Arabidopsis thaliana] 6458 Seq. No. 6851 1.R1040 Contig ID ssr700554846.hl 5'-most EST 6459 Seq. No. 6856 1.R1040 Contig ID $leu7\overline{0}1156170.h1$ 5'-most EST BLASTX Method q1107526 NCBI GI BLAST score 1116 1.0e-122 E value Match length 351 63 % identity (X87931) SIEP1L protein [Beta vulgaris] NCBI Description 6460 Seq. No. 6856 2.R1040 Contig ID q5753379 5'-most EST Method BLASTX g2983289 NCBI GI 154 BLAST score 5.0e-10 E value 73 Match length 42 % identity (AE000703) chaperone DnaJ [Aquifex aeolicus] NCBI Description 6461 Seq. No. Contig ID 6858 1.R1040 5'-most EST epx701104218.hl BLASTX Method g3702339 NCBI GI BLAST score 1046 E value 1.0e-114 340 Match length % identity (AC005397) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 6462 6861 1.R1040 Contig ID LIB3051-049-Q1-K1-B3 5'-most EST BLASTX Method g4164473 NCBI GI BLAST score 367 1.0e-34 E value 105 Match length % identity

(AF061157) negatively light-regulated protein [Vernicia NCBI Description

fordii]

Seq. No.

6463

6861 2.R1040 Contig ID sat701002758.h2 5'-most EST



6464 Seq. No.

6861 3.R1040 Contig ID 5'-most EST qsv701049989.h1

6465 Seq. No.

Contig ID 6862 1.R1040 5'-most EST epx701105792.h1

Method BLASTX q1709358 NCBI GI BLAST score 1224 1.0e-135 E value Match length 411 % identity 57

NCBI Description NUCLEOSIDE-TRIPHOSPHATASE (NUCLEOSIDE TRIPHOSPHATE

PHOSPHOHYDROLASE) (NTPASE) >qi 629638 pir S48859

nucleoside triphosphatase - garden pea

>gi_2129890_pir__S65147 nucleoside triphosphatase

precursor, chromatin-associated - garden pea >gi_563612_emb_CAA83655_ (Z32743) nucleoside triphosphatase

[Pisum satīvum] >gi 4519173 dbj BAA75506.1 (AB022319)

nucleoside triphosphatase (NTPase) [Pisum sativum]

Seq. No. 6466

6862 2.R1040 Contig ID 5'-most EST zsg701126616.hl

Method BLASTX q1709358 NCBI GI BLAST score 750 E value 2.0e-79 Match length 287 % identity 50

NCBI Description NUCLEOSIDE-TRIPHOSPHATASE (NUCLEOSIDE TRIPHOSPHATE

PHOSPHOHYDROLASE) (NTPASE) >gi 629638 pir S48859

nucleoside triphosphatase - garden pea

>gi_2129890_pir__S65147 nucleoside triphosphatase

precursor, chromatin-associated - garden pea

>gi_563612 emb CAA83655 (Z32743) nucleoside triphosphatase

[Pisum sativum] >gi 4519173 dbj BAA75506.1 (AB022319) nucleoside triphosphatase (NTPase) [Pisum sativum]

Seq. No. 6467

6864 1.R1040 Contig ID 5'-most EST pcp700989872.h1

Method BLASTX g3702314 NCBI GI BLAST score 224 9.0e-25 E value Match length 191 46 % identity

NCBI Description (AC002535) similar to SWI/SNF complex subunit BAF170

[Arabidopsis thaliana]

Seq. No. 6468

Contig ID 6864 2.R1040 5'-most EST uC-gmropic011f04b1

Method BLASTX



```
g3702314
NCBI GI
                  426
BLAST score
                  6.0e-42
E value
                  135
Match length
                  60
% identity
                  (AC002535) similar to SWI/SNF complex subunit BAF170
NCBI Description
                   [Arabidopsis thaliana]
                   6469
Seq. No.
                   6866 1.R1040
Contig ID
                  LIB3030-008-Q1-B1-E11
5'-most EST
                  BLASTX
Method
                   g3047116
NCBI GI
BLAST score
                   163
                   3.0e-11
E value
                   38
Match length
% identity
                  (AF058919) No definition line found [Arabidopsis thaliana]
NCBI Description
                   6470
Seq. No.
                   6868 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy126d11b1
                   6471
Seq. No.
                   6869 1.R1040
Contig ID
                   vwf700676052.hl
5'-most EST
                   BLASTX
Method
                   g4191616
NCBI GI
                   635
BLAST score
                   3.0e-66
E value
                   234
Match length
% identity
                   53
                  (AF120334) GTP-binding protein NGB [Homo sapiens]
NCBI Description
                   6472
Seq. No.
                   6871 1.R1040
Contig ID
                   jC-gmfl02220132ab02d1
5'-most EST
                   6473
Seq. No.
                   6872 1.R1040
Contig ID
                   fde700876618.hl
5'-most EST
                   BLASTX
Method
                   q2213584
NCBI GI
                   827
BLAST score
                   3.0e-88
E value
                   295
Match length
                   55
 % identity
NCBI Description (AC000348) T7N9.4 [Arabidopsis thaliana]
                   6474
Seq. No.
                   6874 1.R1040
Contig ID
 5'-most EST
                   k11701208032.h1
                   BLASTX
Method
```

1210

g2330647

4.0e-72

688

288

NCBI GI

E value Match length

BLAST score

W.



% identity (Y14557) ribonucleoprotein [Pisum sativum] NCBI Description

6475 Seq. No.

6874 2.R1040 Contig ID

5'-most EST LIB3167-023-P4-K4-B11

Method BLASTX NCBI GI q2330647 230 BLAST score 7.0e-19 E value 160 Match length 30 % identity

NCBI Description (Y14557) ribonucleoprotein [Pisum sativum]

Seq. No.

6476

6876 1.R1040 Contig ID

5'-most EST LIB3030-008-Q1-B1-F12

Seq. No. 6477

6877 1.R1040 Contig ID

5'-most EST LIB3049-025-Q1-E1-E1

6478 Seq. No.

6877 2.R1040 Contig ID

LIB3030-008-Q1-B1-F2 5'-most EST

BLASTX Method NCBI GI g1749546 245 BLAST score 9.0e-21 E value Match length 89

% identity

(D89169) similar to Saccharomyces cerevisiae SCD6 protein, NCBI Description

SWISS-PROT Accession Number P45978 [Schizosaccharomyces

pombe]

6479 Seq. No.

Contig ID 6877 3.R1040

uC-gmrominsoy116b10b1 5'-most EST

BLASTX Method g3080366 NCBI GI BLAST score 172 3.0e-12E value Match length 60 % identity 53

(AL022580) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No. 6480

6877 4.R1040 Contig ID

5'-most EST LIB3\overline{1}07-014-Q1-K1-C9

Seq. No. 6481

Contig ID 6878 1.R1040

LIB3107-009-Q1-K1-A3 5'-most EST

Method BLASTX q1619602 NCBI GI BLAST score 503 7.0e-51 E value

1211



Match length 147 % identity 65

NCBI Description (Y08726) MtN3 [Medicago truncatula]

Seq. No. 6482

Contig ID 6878_3.R1040 5'-most EST kl1701206730.h1

Method BLASTX
NCBI GI g1619602
BLAST score 272
E value 2.0e-24
Match length 63
% identity 81

NCBI Description (Y08726) MtN3 [Medicago truncatula]

Seq. No. 6483

Contig ID 6880_1.R1040

5'-most EST LIB3051-022-Q1-K1-D4

Method BLASTX
NCBI GI g1708975
BLAST score 448
E value 7.0e-44
Match length 396
% identity 35

NCBI Description METHIONINE GAMMA-LYASE (L-METHIONINASE)

>gi_1084040_pir__JC4174 methionine gamma-lyase (EC
4.4.1.11) - Pseudomonas putida >gi_902932_dbj_BAA13642_
(D88554) L-methionine gamma-lyase [Pseudomonas putida]
>gi_2217945_dbj_BAA20553_ (D89015) L-methionine gamma-lyase

[Pseudomonas putida] >gi_1096717_prf__2112270A Met

gamma-lyase [Pseudomonas putida]

Seq. No. 6484

Contig ID 6880 2.R1040 5'-most EST g5606623 Method BLASTX NCBI GI g1709007 BLAST score 172 E value 4.0e-12 Match length 102 % identity 38

NCBI Description O-SUCCINYLHOMOSERINE SULFHYDRYLASE (OSH SULFHYDRYLASE)

>gi_607830 (U10904) O-succinylhomoserine sulfhydrylase

[Pseudomonas aeruginosa]

Seq. No. 6485

Contig ID 6880 3.R1040

5'-most EST LIB3065-005-Q1-N1-A2

Seq. No. 6486

Contig ID 6881_1.R1040

5'-most EST LIB3056-004-Q1-N1-D10

Method BLASTX
NCBI GI 9481815
BLAST score 195
E value 1.0e-14
Match length 188



% identity 30

NCBI Description vegetative storage protein - western balsam poplar x

cottonwood >gi 309839 (L20233) vegetative storage protein

[Populus trichocarpa x Populus deltoides]

Seq. No. 6487

Contig ID 6884 1.R1040

5'-most EST uC-gmrominsoy103a04b1

Method BLASTX
NCBI GI 94415937
BLAST score 160
E value 8.0e-11
Match length 133
% identity 28

NCBI Description (AC006418) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6488

Contig ID 6886 1.R1040

5'-most EST LIB3030-008-Q1-B1-D10

Method BLASTN
NCBI GI g402752
BLAST score 219
E value 1.0e-120
Match length 419
% identity 96

NCBI Description G.max gene fusA

Seq. No. 6489

6887 1.R1040 Contig ID 5'-most EST q5688103 Method BLASTX NCBI GI q3549669 BLAST score 640 8.0e-67 E value Match length 216 % identity 42

NCBI Description (ALO31394) putative protein [Arabidopsis thaliana]

Seq. No. 6490

Contig ID 6888_1.R1040 5'-most EST hyd700731116.h1

Method BLASTX
NCBI GI g3152666
BLAST score 221
E value 9.0e-18
Match length 111
% identity 38

NCBI Description (AF064606) KB07 protein [Homo sapiens]

Seq. No. 6491

Contig ID 6889 1.R1040

5'-most EST LIB3030-008-Q1-B1-D3

Seq. No. 6492

Contig ID 6891 1.R1040

5'-most EST LIB3051-013-Q1-E1-E11

Method BLASTX



NCBI GI q2052508 BLAST score 917 E value 1.0e-120 Match length 360 % identity 61

NCBI Description (U95759) lysine-ketoglutarate reductase/saccharopine dehydrogenase bifunctional enzyme [Arabidopsis thaliana]

Seq. No. 6493

6891 2.R1040 Contig ID zhf700963155.h1 5'-most EST

Method BLASTX NCBI GI g2052510 BLAST score 858 E value 2.0e-92 232 Match length % identity 72

NCBI Description (U95758) lysine-ketoglutarate reductase/saccharopine dehydrogenase bifunctional enzyme [Arabidopsis thaliana]

Seq. No.

6494 Contig ID 6894 1.R1040 5'-most EST vwf700674245.h1

Seq. No.

6495 6894 2.R1040

Contig ID 5'-most EST

fde700873040.h1

Seq. No.

6496

Contig ID 5'-most EST 6897 1.R1040 jex700905170.hl

Method BLASTX g2979553 NCBI GI BLAST score 586 3.0e-60 E value Match length 291 % identity

NCBI Description (AC003680) hypothetical protein [Arabidopsis thaliana]

6497 Seq. No.

Contig ID 6899 1.R1040

5'-most EST uC-gmflminsoy008e10b1

Method BLASTX NCBI GI g416758 BLAST score 338 E value 3.0e-35 Match length 124 % identity 59

NCBI Description SERINE CARBOXYPEPTIDASE PRECURSOR >qi 166674 (M81130)

carboxypeptidase Y-like protein [Arabidopsis thaliana] >gi_445120_prf__1908426A carboxypeptidase Y [Arabidopsis

thaliana]

6498 Seq. No.

6901 1.R1040 Contig ID 5'-most EST zhf700952535.h1

Method BLASTN



NCBI GI g169127 BLAST score 209 E value 1.0e-113 Match length 561 % identity 84

NCBI Description Pisum sativum (clone pCLp) nuclear encoded precursor to

chloroplast protein mRNA, complete cds

Seq. No. 6499

Contig ID 6901_2.R1040 5'-most EST awf700839723.h1

Method BLASTX
NCBI GI g461753
BLAST score 307
E value 6.0e-28
Match length 127
% identity 45

NCBI Description ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPA HOMOLOG

PRECURSOR >gi_419773_pir__S31164 ATP-dependent ClpB

proteinase regulatory chain homolog precursor, chloroplast - garden pea >gi_169128 (L09547) nuclear encoded precursor

to chloroplast protein [Pisum sativum]

Seq. No. 6500

Contig ID 6907_1.R1040

5'-most EST LIB3030-008-Q1-B1-C10

Method BLASTX
NCBI GI g1731141
BLAST score 165
E value 3.0e-11
Match length 153
% identity 38

NCBI Description HYPOTHETICAL 90.8 KD PROTEIN T05H10.7 IN CHROMOSOME II

>gi 3878112 emb CAA87788_ (Z47811) similar to

glycerophosphoryl diester phosphodiesterase domain; cDNA EST EMBL:D27842 comes from this gene; cDNA EST EMBL:D27841 comes from this gene; cDNA EST EMBL:D32990 comes from this

gene; cDNA EST EMBL:D35712 comes...

>qi 3879502 emb CAA87796_ (Z47812) similar to

glycerophosphoryl diester phosphodiesterase domain; cDNA EST EMBL:D27842 comes from this gene; cDNA EST EMBL:D27841 comes from this gene; cDNA EST EMBL:D32990 comes from this

gene; cDNA EST EMBL: D35712 comes

Seq. No. 6501

Contig ID 6916_1.R1040
5'-most EST g4293044
Method BLASTX
NCBI GI g3287681
BLAST score 468
E value 3.0e-96

Match length 304 % identity 66

NCBI Description (AC003979) ESTs gb_Z34075, gb_Z34835 and gb_AA404888 come

from this gene. [Arabidopsis thaliana]

Seq. No. 6502



```
6920 1.R1040
Contig ID
                   ssr700555158.hl
5'-most EST
                   6503
Seq. No.
                   6922 1.R1040
Contig ID
                   LIB3030-008-Q1-B1-A3
5'-most EST
                   BLASTX
Method
                   g3135254
NCBI GI
                   415
BLAST score
                   9.0e-41
E value
                   127
Match length
                   62
% identity
NCBI Description (AC003058) hypothetical protein [Arabidopsis thaliana]
                   6504
Seq. No.
                   6923 1.R1040
Contig ID
5'-most EST
                   zsq701122147.hl
                   6505
Seq. No.
                   6923 2.R1040
Contig ID
                   LIB3109-046-Q1-K1-F8
5'-most EST
                   6506
Seq. No.
                   6925 1.R1040
Contig ID
                   jC-g\overline{m}st02400061e07a1
5'-most EST
                   BLASTX
Method
                   g2982266
NCBI GI
BLAST score
                   437
                   3.0e-43
E value
                   101
Match length
% identity
                   (AF051216) probable fibrillarin [Picea mariana]
NCBI Description
                   6507
Seq. No.
                    6925 2.R1040
Contig ID
                   gsv701045236.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                    g2982266
BLAST score
                    1044
                    1.0e-114
E value
Match length
                    244
                    80
% identity
                   (AF051216) probable fibrillarin [Picea mariana]
NCBI Description
                    6508
Seq. No.
                    6925 3.R1040
Contig ID
                    ssr7\overline{0}0557074.h1
5'-most EST
                    BLASTN
Method
                    g21002
NCBI GI
                    80
BLAST score
                    1.0e-36
E value
Match length
                    226
 % identity
                    11
                    Bean DNA for glycine-rich cell wall protein GRP 1.8
NCBI Description
```

1216

6509

6925_4.R1040

Seq. No.

Contig ID



5'-most EST LIB3051-067-Q1-K1-A11

Method BLASTX
NCBI GI g2982266
BLAST score 452
E value 5.0e-45
Match length 104
% identity 82

NCBI Description (AF051216) probable fibrillarin [Picea mariana]

Seq. No. 6510

Contig ID 6925_5.R1040 5'-most EST vzy700754605.h1

Method BLASTX
NCBI GI g2982266
BLAST score 433
E value 1.0e-42
Match length 97
% identity 82

NCBI Description (AF051216) probable fibrillarin [Picea mariana]

Seq. No. 6511

Contig ID 6925 6.R1040

5'-most EST jC-gmst02400009b07a1

Seq. No. 6512

Contig ID 6925_7.R1040 5'-most EST jsh701067142.h1

Method BLASTN
NCBI GI g21002
BLAST score 84
E value 1.0e-39
Match length 206
% identity 9

NCBI Description Bean DNA for glycine-rich cell wall protein GRP 1.8

Seq. No. 6513

Contig ID 6926_1.R1040 5'-most EST rlr700897678.h1

Seq. No. 6514

6933 1.R1040 Contig ID 5'-most EST g4305391 Method BLASTX NCBI GI g2576411 BLAST score 745 5.0e-85 E value Match length 226 % identity 74

NCBI Description (AF012833) similar to dynamin-like protein encoded by GenBank Accession Number X99669 [Arabidopsis thaliana]

Seq. No. 6515

Contig ID 6933_2.R1040 5'-most EST leu701152583.h1

Method BLASTX NCBI GI g2576411 BLAST score 221



E value 6.0e-18
Match length 71
% identity 63

NCBI Description (AF012833) similar to dynamin-like protein encoded by GenBank Accession Number X99669 [Arabidopsis thaliana]

Seq. No. 6516

Contig ID 6933_3.R1040 5'-most EST zhf700955208.h1

Method BLASTX
NCBI GI g2576411
BLAST score 329
E value 1.0e-30
Match length 101
% identity 67

NCBI Description (AF012833) similar to dynamin-like protein encoded by GenBank Accession Number X99669 [Arabidopsis thaliana]

Seq. No. 6517

Contig ID 6934_1.R1040

5'-most EST LIB3030-007-Q1-B1-G6

Method BLASTX
NCBI GI g1707015
BLAST score 171
E value 5.0e-12
Match length 131
% identity 32

NCBI Description (U78721) protein phosphatase 2C isolog [Arabidopsis

thaliana]

Seq. No. 6518

Contig ID 6935_1.R1040 5'-most EST zhf700963017.h1

Method BLASTX
NCBI GI g4468812
BLAST score 808
E value 2.0e-86
Match length 221
% identity 66

NCBI Description (AL035601) putative protein [Arabidopsis thaliana]

Seq. No. 6519

Contig ID 6937_1.R1040

5'-most EST LIB3109-031-Q1-K1-D10

Method BLASTX
NCBI GI g1351651
BLAST score 215
E value 6.0e-17
Match length 163
% identity 31

NCBI Description HYPOTHETICAL 43.7 KD PROTEIN C24B11.08C IN CHROMOSOME I

>gi_2130353_pir__S62553 hypothetical protein SPAC24B11.08c

- fission yeast (Schizosaccharomyces pombe) >gi_1061296_emb_CAA91773_ (Z67757) unknown

[Schizosaccharomyces pombe]

Seq. No. 6520



Contig ID 6938_1.R1040 5'-most EST zhf700963141.h1 Method BLASTN

Method BLASTN
NCBI GI g1853969
BLAST score 78
E value 7.0e-36
Match length 283
% identity 82

NCBI Description Vigna unguiculata mRNA for CPRD46 protein, complete cds

Seq. No. 6521

Contig ID 6945_1.R1040

5'-most EST uC-gmropic055e06b1

Seq. No. 6522

Contig ID 6947_1.R1040 5'-most EST zhf700957949.h1

Seq. No. 6523

Contig ID 6948_1.R1040 5'-most EST kl1701206149.h1

Method BLASTX
NCBI GI g2281627
BLAST score 329
E value 8.0e-36
Match length 138
% identity 60

NCBI Description (AF003094) AP2 domain containing protein RAP2.1

[Arabidopsis thaliana]

Seq. No. 6524

Contig ID 6948_2.R1040 5'-most EST trc700565376.h1

Seq. No. 6525

Contig ID 6949_1.R1040 5'-most EST txt700733044.h1

Method BLASTX
NCBI GI g3393062
BLAST score 620
E value 3.0e-64
Match length 208
% identity 55

NCBI Description (Y17386) putative In2.1 protein [Triticum aestivum]

Seq. No. 6526

Contig ID 6949_2.R1040

5'-most EST uC-gmronoir045f04b1

Seq. No. 6527

Contig ID 6951 1.R1040 5'-most EST vzy700751782.h1

Method BLASTX
NCBI GI g2351380
BLAST score 526
E value 3.0e-53
Match length 285



```
% identity 41

NCBI Description (U54559) translation initiation factor eIF3 p40 subunit [Homo sapiens] >gi_4503515_ref_NP_003747.1_pEIF3S3_UNKNOWN

Seq. No. 6528

Contig ID 6951_2.R1040
5'-most EST kmv700738117.h1

Method BLASTX
```

Method BLASTX
NCBI GI g2351380
BLAST score 133
E value 7.0e-13
Match length 149
% identity 36

NCBI Description (U54559) translation initiation factor eIF3 p40 subunit [Homo sapiens] >gi_4503515_ref_NP_003747.1_pEIF3S3_ UNKNOWN

Seq. No. 6529
Contig ID 6952_1.R1040
5'-most EST LIB3030-006-Q1-B1-E12
Mathod BLASTY

Method BLASTX
NCBI GI g1762584
BLAST score 269
E value 2.0e-23
Match length 99
% identity 49

NCBI Description (U63373) polygalacturonase isoenzyme 1 beta subunit homolog [Arabidopsis thaliana]

Seq. No. 6530
Contig ID 6952_2.R1040
5'-most EST uC-gmropic049h03b1

Method BLASTX NCBI GI g1762584 BLAST score 187

E value 6.0e-14
Match length 59
% identity 59

NCBI Description (U63373) polygalacturonase isoenzyme 1 beta subunit homolog [Arabidopsis thaliana]

Seq. No. 6531 Contig ID 6956_1.R1040

5'-most EST LIB3170-021-Q1-K1-A12

Seq. No. 6532

Contig ID 6960 1.R1040

5'-most EST uC-gmrominsoy066e08b1

Method BLASTX
NCBI GI g3182915
BLAST score 630
E value 2.0e-65
Match length 180
% identity 62

NCBI Description ADP-RIBOSYLATION FACTOR >gi_1842150_dbj_BAA19225_ (AB001051) ADP-ribosylation factor [Dugesia japonica]

Seq. No. 6533



Contig ID 6960 2.R1040 5'-most EST asn701134890.h2

Method BLASTX NCBI GI q461532 BLAST score 432 E value 1.0e-42 Match length 113 % identity 66

ADP-RIBOSYLATION FACTOR >gi_1362500 pir D49993 NCBI Description

ADP-ribosylation factor - Ajellomyces capsulata >gi_407693 (L25117) ADP-ribosylation factor [Histoplasma capsulatum]

Seq. No. 6534

Contig ID 6963 1.R1040

5'-most EST jC-gmro02910051g04a1

Method BLASTX NCBI GI q2829918 BLAST score 1073 1.0e-117 E value Match length 309 72 % identity

NCBI Description (AC002291) similar to "tub" protein gp U82468 2072162

[Arabidopsis thaliana]

Seq. No. 6535

Contig ID 6963 2.R1040

5'-most EST LIB3052-013-Q1-N1-C6

Method BLASTX NCBI GI q2829918 BLAST score 472 5.0e-47 E value Match length 176 67

% identity

NCBI Description (AC002291) similar to "tub" protein gp_U82468_2072162

[Arabidopsis thaliana]

Seq. No. 6536

Contig ID 6963 3.R1040 5'-most EST dpv701103595.h1

Method BLASTX q2829918 NCBI GI BLAST score 335 E value 2.0e-31 Match length 95 % identity 71

(AC002291) similar to "tub" protein gp_U82468_2072162 NCBI Description

[Arabidopsis thaliana]

Seq. No. 6537

Contig ID 6963 5.R1040

5'-most EST LIB3170-032-Q1-K1-G11

Method BLASTX NCBI GI q2829918 BLAST score 191 E value 2.0e-14 Match length 69 % identity 64



```
(AC002291) similar to "tub" protein gp_U82468_2072162
NCBI Description
                  [Arabidopsis thaliana]
                  6538
Seq. No.
                  6963 6.R1040
Contig ID
                  zzp700831683.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2829918
                  226
BLAST score
                  2.0e-18
E value
                  70
Match length
                  73
% identity
                   (AC002291) similar to "tub" protein gp_U82468_2072162
NCBI Description
                   [Arabidopsis thaliana]
                  6539
Seq. No.
                  6963 7.R1040
Contig ID
                  awf700841234.h1
5'-most EST
                  BLASTX
Method
                  g2829918
NCBI GI
                  219
BLAST score
                   7.0e-24
E value
                  80
Match length
                   81
% identity
                   (AC002291) similar to "tub" protein gp_U82468_2072162
NCBI Description
                   [Arabidopsis thaliana]
                   6540
Seq. No.
                   6965 1.R1040
Contig ID
                   LIB3049-055-Q1-E1-G12
5'-most EST
                   BLASTX
Method
                   g3335347
NCBI GI
                   481
BLAST score
                   5.0e-48
E value
                   163
Match length
% identity
                   (AC004512) Contains similarity to ARI, RING finger protein
NCBI Description
                   gb X98309 from Drosophila melanogaster. ESTs gb_T44383,
                   gb_W43120, gb_N65868, gb_H36013, gb_AA042241, gb_T76869 and
                   gb_AA042359 come from this gene. [Arabidopsis thaliana]
 Seq. No.
                   6965 2.R1040
 Contig ID
                   sat701007023.h1
 5'-most EST
                   BLASTX
 Method
                   g3335347
 NCBI GI
 BLAST score
                   1324
                   1.0e-147
 E value
 Match length
                   272
                   83
 % identity
                   (AC004512) Contains similarity to ARI, RING finger protein
 NCBI Description
                   gb X98309 from Drosophila melanogaster. ESTs gb T44383,
                   gb_W43120, gb_N65868, gb_H36013, gb_AA042241, gb_T76869 and
                   gb_AA042359 come from this gene. [Arabidopsis thaliana]
```

Seq. No. 6542

Contig ID 6965_3.R1040



5'-most EST LIB3051-106-Q1-K1-A7

Method BLASTX NCBI GI g3335347 BLAST score 819 1.0e-87 E value 250 Match length 61 % identity

(AC004512) Contains similarity to ARI, RING finger protein NCBI Description

gb X98309 from Drosophila melanogaster. ESTs gb_T44383, gb_W43120, gb_N65868, gb_H36013, gb_AA042241, gb_T76869 and

gb_AA042359 come from this gene. [Arabidopsis thaliana]

Seq. No. 6543

Contig ID 6965 4.R1040 5'-most EST fde700877103.h1

Method BLASTX NCBI GI g3335347 BLAST score 180 E value 4.0e-13 Match length 92 % identity 45

(AC004512) Contains similarity to ARI, RING finger protein NCBI Description

gb X98309 from Drosophila melanogaster. ESTs gb T44383, gb_W43120, gb_N65868, gb_H36013, gb_AA042241, gb_T76869 and gb AA042359 come from this gene. [Arabidopsis thaliana]

Seq. No. 6544

Contig ID 6966 1.R1040

5'-most EST jC-qmf102220089q12a1

6545 Seq. No.

Contig ID 6968 1.R1040

5'-most EST LIB3040-043-Q1-E1-A10

Method BLASTX NCBI GI g2129726 BLAST score 953 E value 1.0e-103 Match length 226

% identity

NCBI Description RNA polymerase II third largest chain RPB35.5A -

Arabidopsis thaliana >gi 514318 (L34770) RNA polymerase II

third largest subunit [Arabidopsis thaliana]

>gi 4544370 gb AAD22281.1 AC006920 5 (AC006920) RNA

polymerase II, third largest subunit [Arabidopsis thaliana]

Seq. No. 6546

Contig ID 6972 1.R1040 leu701151137.h1 5'-most EST

BLASTX Method NCBI GI g2281086 BLAST score 361 E value 4.0e-34 Match length 195 % identity 38

(AC002333) indole-3-acetate beta-glucosyltransferase isolog NCBI Description

[Arabidopsis thaliana]



```
6547
Seq. No.
                  6974_1.R1040
Contig ID
                  jC-gmst02400009f05a1
5'-most EST
                  BLASTX
Method
                  g3600054
NCBI GI
BLAST score
                  221
                  6.0e-18
E value
                  129
Match length
% identity
                 (AF080120) No definition line found [Arabidopsis thaliana]
NCBI Description
                  6548
Seq. No.
                  6976 1.R1040
Contig ID
                  bth700844626.hl
5'-most EST
                  BLASTX
Method
                  g4115931
NCBI GI
                  472
BLAST score
                  3.0e-47
E value
                  182
Match length
                   58
% identity
                   (AF118223) contains similarity to Guillardia theta ABC
NCBI Description
                  transporter (GB:AF041468) [Arabidopsis thaliana]
                   6549
Seq. No.
                   6976 2.R1040
Contig ID
                  LIB3051-059-Q1-K2-G7
5'-most EST
                   BLASTX
Method
                   g4115931
NCBI GI
                   371
BLAST score
                   2.0e-35
E value
Match length
                   149
% identity
                   (AF118223) contains similarity to Guillardia theta ABC
NCBI Description
                   transporter (GB:AF041468) [Arabidopsis thaliana]
                   6550
Seq. No.
                   6977 1.R1040
Contig ID
5'-most EST
                   zzp700832937.h1
Method
                   BLASTX
                   q2980770
NCBI GI
                   1214
BLAST score
E value
                   1.0e-134
Match length.
                   323
% identity
                   (AL022198) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   6551
Seq. No.
                   6977 2.R1040
Contig ID
                   uC-gmrominsoy080d04b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2980770
BLAST score
                   398
E value
                   1.0e-38
```

(AL022198) putative protein kinase [Arabidopsis thaliana]

120

64

Match length % identity

NCBI Description



```
6552
Seq. No.
                   6982 1.R1040
Contig ID
                   LIB3049-044-Q1-E1-G2
5'-most EST
                   6553
Seq. No.
                   6983 1.R1040
Contig ID
                   gsv701053487.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2275210
BLAST score
                   454
                   5.0e-45
E value
                   175
Match length
                   51
% identity
                   (AC002337) peptidyl-prolyl cis-trans isomerase isolog
NCBI Description
                   [Arabidopsis thaliana]
                   6554
Seq. No.
                   6983 2.R1040
Contig ID
                   fua701043243.h1
5'-most EST
                   BLASTX
Method
                   g2275210
NCBI GI
BLAST score
                   194
                   8.0e-33
E value
Match length
                   174
                   50
% identity
                   (AC002337) peptidyl-prolyl cis-trans isomerase isolog
NCBI Description
                   [Arabidopsis thaliana]
                   6555
Seq. No.
                   6985 1.R1040
Contig ID
                   uaw700665809.hl
5'-most EST
                   6556
Seq. No.
                   6987 1.R1040
Contig ID
                   g4395747
5'-most EST
                   BLASTN
Method
                   g19655
NCBI GI
                    466
BLAST score
                    0.0e + 00
E value
                    510
Match length
                    98
 % identity
                   M.sativa 26S rRNA
NCBI Description
                    6557
 Seq. No.
                    6987 2.R1040
 Contig ID
                    zpv700760254.h1
 5'-most EST
                    BLASTN
Method
                    q311302
 NCBI GI
 BLAST score
                    42
                    1.0e-13
 E value
                    102
 Match length
                    85
 % identity
                   N.sylvestris mRNA
 NCBI Description
```

5'-most EST asn701140057.h1

6987 3.R1040

6558

Seq. No.

Contig ID



BLASTN Method g311302 NCBI GI 33 BLAST score 6.0e-09 E value 89 Match length 84 % identity NCBI Description N.sylvestris mRNA 6559 Seq. No. 6987 4.R1040 Contig ID LIB3107-029-Q1-K1-B2 5'-most EST BLASTN Method q22553 NCBI GI 76 BLAST score 1.0e-34 E value 80 Match length 99 % identity Tomato 25S ribosomal RNA gene NCBI Description 6560 Seq. No. 6989 1.R1040 Contig ID rlr700896221.hl 5'-most EST 6561 Seq. No. 6991 1.R1040 Contig ID jC-gmst02400049e03d1 5'-most EST 6562 Seq. No. 6994 1.R1040 Contig ID LIB3030-007-Q1-B1-A1 5'-most EST BLASTN Method g3290208 NCBI GI 51 BLAST score 1.0e-19 E value 71 Match length 93 % identity NCBI Description Malus domestica MADS-box protein 1 mRNA, complete cds Seq. No. 6563 6999 1.R1040 Contig ID LIB3051-025-Q1-K1-F2 5'-most EST 6564 Seq. No. Contig ID 6999 2.R1040 LIB3107-015-Q1-K1-A2 5'-most EST 6565 Seq. No.

Contig ID 7001_1.R1040

5'-most EST uC-gmropic043c05b1

Method BLASTX
NCBI GI g3482933
BLAST score 1260
E value 1.0e-139
Match length 552

% identity 49 NCBI Description (AC003970) Similar to cdc2 protein kinases [Arabidopsis

thaliana]

% identity

NCBI Description

70



```
6566
Seg. No.
                  7001 2.R1040
Contig ID
                  fC-gmle700873234y1
5'-most EST
                  BLASTX
Method
                  g3482933
NCBI GI
                  691
BLAST score
                   2.0e-72
E value
Match length
                  395
                   42
% identity
                   (AC003970) Similar to cdc2 protein kinases [Arabidopsis
NCBI Description
                   thaliana]
                   6567
Seq. No.
                   7001 5.R1040
Contig ID
                   rca700999051.hl
5'-most EST
                   6568
Seq. No.
                   7001 6.R1040
Contig ID
                   fC-gmro700870190d1
5'-most EST
                   6569
Seq. No.
                   7010 1.R1040
Contig ID
                   LIB3030-006-Q1-B1-G3
5'-most EST
Method
                   BLASTX
                   g2191161
NCBI GI
BLAST score
                   760
                   7.0e-81
E value
                   218
Match length
                   68
% identity
                   (AF007270) contains similarity to B. subtilus flagellar
NCBI Description
                   biosynthesis protein FLHA (SW:P35620) [Arabidopsis
                   thaliana]
                   6570
Seq. No.
                   7011 1.R1040
Contig ID
                   LIB3109-001-Q1-K2-A9
5'-most EST
                   BLASTX
Method
                   g4206122
NCBI GI
                   1068
BLAST score
                   1.0e-128
E value
                   283
Match length
                   80
 % identity
                   (AF097667) protein phosphatase 2C homolog [Mesembryanthemum
NCBI Description
                   crystallinum]
                   6571
 Seq. No.
                   7012 1.R1040
 Contig ID
                    fua701040701.hl
 5'-most EST
                   BLASTX
Method
                   q1621268
 NCBI GI
                   1243
 BLAST score
                    1.0e-146
 E value
 Match length
                    364
```

(Z81012) unknown [Ricinus communis]



```
6572
Seq. No.
                  7016 1.R1040
Contig ID
                  epx701109085.hl
5'-most EST
                  BLASTX
Method
                  g2827655
NCBI GI
                  270
BLAST score
                  1.0e-23
E value
                  109
Match length
                   51
% identity
                 (AL021637) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7017 1.R1040
Contig ID
                  pmv700890924.hl
5'-most EST
                   6574
Seq. No.
                   7017 2.R1040
Contig ID
                   jsh701070327.h1
5'-most EST
Seq. No.
                   7017 3.R1040
Contig ID
                   wvk7\overline{0}0685462.h1
5'-most EST
                   6576
Seq. No.
                   7020 1.R1040
Contig ID
                   LIB3106-034-Q1-K1-D2
5'-most EST
                   BLASTX
Method
                   g2506211
NCBI GI
                   1679
BLAST score
                   0.0e+00
E value
                   333
Match length
                   97
% identity
                   VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A (V-ATPASE 69 KD
NCBI Description
                   SUBUNIT) (VAA3-1) >gi 849136 (U26709) vacuolar H+-ATPase
                   subunit A [Vigna radiata]
                   6577
Seq. No.
                   7020 2.R1040
Contig ID
                   LIB3139-035-P1-N1-B8
5'-most EST
Method
                   BLASTX
                   q3169287
NCBI GI
BLAST score
                   1893
E value
                   0.0e + 00
Match length
                   387
 % identity
                   (AF050673) vacuolar H+-ATPase catalytic subunit [Gossypium
NCBI Description
                   hirsutum]
                    6578
 Seq. No.
                   7020 4.R1040
 Contig ID
                   LIB3170-084-Q1-K1-A5
 5'-most EST
                   BLASTN
Method
 NCBI GI
                   q849135
 BLAST score
                    83
```

1228

9.0e-39

134

96

E value

Match length

% identity



NCBI Description Vigna radiata vacuolar H+-ATPase subunit A mRNA, complete cds

Seq. No. 6579

Contig ID 7020_6.R1040 5'-most EST jex700905426.h1

Method BLASTN
NCBI GI 9849135
BLAST score 140
E value 5.0e-73
Match length 180
% identity 94

NCBI Description Vigna radiata vacuolar H+-ATPase subunit A mRNA, complete

cds

Seq. No. 6580

Contig ID 7022_1.R1040

5'-most EST LIB3092-021-Q1-K1-D2

Method BLASTX
NCBI GI g4079809
BLAST score 253
E value 3.0e-21
Match length 198
% identity 3

NCBI Description (AF071172) HERC2 [Homo sapiens]

Seq. No. 6581

Contig ID 7023_1.R1040 5'-most EST jex700904926.h1

Method BLASTX
NCBI GI g3080389
BLAST score 375
E value 3.0e-72
Match length 187
% identity 74

NCBI Description (AL022603) putative membrane associated protein

[Arabidopsis thaliana]

Seq. No. 6582

Contig ID 7024_1.R1040 5'-most EST pmv700894358.h1

Seq. No. 6583

Contig ID 7025_1.R1040 5'-most EST leu701145072.h1

Method BLASTX
NCBI GI g4455313
BLAST score 469
E value 6.0e-47
Match length 155
% identity 54

NCBI Description (AL035528) fatty acid elongase-like protein (cer2-like)

[Arabidopsis thaliana]

Seq. No. 6584

Contig ID 7025_2.R1040

5'-most EST LIB3055-005-Q1-N1-C10



```
BLASTX
Method
                  q4455313
NCBI GI
                  418
BLAST score
                  5.0e-41
E value
                  153
Match length
                  50
% identity
                  (AL035528) fatty acid elongase-like protein (cer2-like)
NCBI Description
                   [Arabidopsis thaliana]
                   6585
Seq. No.
                  7029 1.R1040
Contig ID
                  LIB3051-026-Q1-K1-F1
5'-most EST
                  BLASTX
Method
                   q4510385
NCBI GI
                   186
BLAST score
                   1.0e-13
E value
                   48
Match length
                   75
% identity
NCBI Description (AC007017) unknown protein [Arabidopsis thaliana]
                   6586
Seq. No.
                   7029 2.R1040
Contig ID
5'-most EST
                   hyd700726122.hl
                   6587
Seq. No.
                   7029 3.R1040
Contig ID
5'-most EST
                   k117\overline{0}1212057.h1
                   6588
Seq. No.
                   7029 4.R1040
Contig ID
                   zzp700835108.h1
5'-most EST
                   6589
Seq. No.
                   7031 1.R1040
Contig ID
                   LIB3030-006-Q1-B1-F1
5'-most EST
                   BLASTX
Method
                   g3834314
NCBI GI
BLAST score
                   588
                   6.0e-61
E value
                   159
Match length
                   75
 % identity
                   (AC005679) Similar to gene pi010 glycosyltransferase
NCBI Description
                   gi_2257490 from S. pombe clone 1750 gb_AB004534. ESTs
                   gb_T46079 and gb_AA394466 come from this gene. [Arabidopsis
                   thaliana]
                    6590
 Seq. No.
                    7033 1.R1040
 Contig ID
                   LIB3107-038-Q1-K1-H5
 5'-most EST
                    6591
 Seq. No.
```

7037 1.R1040 Contig ID $vzy7\overline{0}0750957.h1$ 5'-most EST

BLASTX Method g3080416 NCBI GI 237 BLAST score 6.0e-39 E value



Match length 102 % identity 86 NCBI Description (AL [Ar

(AL022604) UDP-galactose transporter - like protein

[Arabidopsis thaliana]

Seq. No. 6592

Contig ID 7039_1.R1040 5'-most EST zsg701129960.h1

Method BLASTX
NCBI GI g3213227
BLAST score 167
E value 2.0e-11
Match length 90
% identity 34

NCBI Description (AF035209) putative v-SNARE Vtila [Mus musculus]

>gi_3421062 (AF035823) 29-kDa Golgi SNARE [Mus musculus]

Seq. No. 6593

Contig ID 7041_1.R1040

5'-most EST LIB3049-026-Q1-E1-F1

Seq. No. 6594

Contig ID 7046_1.R1040

5'-most EST jC-gmst02400030d06a1

Method BLASTX
NCBI GI g2811122
BLAST score 267
E value 4.0e-23
Match length 152
% identity 40

NCBI Description (U87318) NaDC-2 [Xenopus laevis]

Seq. No. 6595

Contig ID 7047_1.R1040 5'-most EST jex700909603.h1

Seq. No. 6596

Contig ID 7056_1.R1040 5'-most EST uC-gmropic023e04b1

Method BLASTX
NCBI GI 94455210
BLAST score 138
E value 6.0e-09
Match length 112

% identity 46 NCBI Description (AL035440) putative aspartate-tRNA ligase [Arabidopsis

thaliana]

Seq. No. 6597

Contig ID 7058_1.R1040 5'-most EST g4290079

Seq. No. 6598

Contig ID 7058 2.R1040

5'-most EST jC-gmf102220076c05a1

Seq. No. 6599

Seq. No.

Contig ID

6606

7062 2.R1040

```
Contig ID
                   7060 1.R1040
5'-most EST
                   LIB3139-118-P1-N1-E3
Method
                   BLASTX
NCBI GI
                   g1737222
BLAST score
                   2812
E value
                   0.0e + 00
Match length
                   583
% identity
                   86
                  (U79958) BP-80 vacuolar sorting receptor [Pisum sativum]
NCBI Description
                   6600
Seq. No.
                   7060 2.R1040
Contig ID
5'-most EST
                   zhf7\overline{0}0953252.h1
                   BLASTX
Method
NCBI GI
                   g1737222
BLAST score
                   461
E value
                   9.0e-46
Match length
                   105
% identity
                   84
                  (U79958) BP-80 vacuolar sorting receptor [Pisum sativum]
NCBI Description
Seq. No.
                   7060 3.R1040
Contig ID
5'-most EST
                   LIB3040-016-Q1-E1-C12
Method
                   BLASTX
NCBI GI
                   g2911587
BLAST score
                   436
E value
                   5.0e-43
Match length
                   160
% identity
                   54
NCBI Description
                   (AJ224442) methyltransferase [Homo sapiens]
                   6602
Seq. No.
Contig ID
                   7060 4.R1040
5'-most EST
                   djj700605996.h2
Method
                   BLASTN
NCBI GI
                   q414831
BLAST score
                   67
E value
                   5.0e-29
                   71
Match length
% identity
                   99
NCBI Description
                  Glycine max (Rablp) mRNA, complete cds
Seq. No.
                   6603
                   7060 5.R1040
Contig ID
5'-most EST
                   LIB3109-032-Q1-K1-C2
                   6604
Seq. No.
                   7061 1.R1040
Contig ID
5'-most EST
                   LIB3030-006-Q1-B1-B11
Seq. No.
                   6605
Contig ID
                   7062 1.R1040
5'-most EST
                   uC-qmrominsoy060c08b1
```



5'-most EST LIB3030-006-Q1-B1-B12

Method BLASTX
NCBI GI g2500583
BLAST score 148
E value 5.0e-09
Match length 101
% identity 36

NCBI Description RNA-BINDING PROTEIN EWS >gi_1083465_pir__A55726 RNA-binding

protein Ews - mouse >gi 488513 emb CAA55815 (X79233) EWS

[Mus musculus]

Seq. No. 6607

Contig ID 7064 1.R1040

5'-most EST LIB3139-107-P1-N1-B4

Method BLASTX
NCBI GI g3785990
BLAST score 510
E value 2.0e-51
Match length 172
% identity 60

NCBI Description (AC005560) hypothetical protein [Arabidopsis thaliana]

>gi 4220487 (AC006069) hypothetical protein [Arabidopsis

thaliana]

Seq. No. 6608

7068 1.R1040 Contig ID 5'-most EST q4397127 BLASTX Method NCBI GI q2352492 BLAST score 571 1.0e-58 E value Match length 237 % identity 49

NCBI Description (AF005047) transport inhibitor response 1 [Arabidopsis

thaliana] >gi_2352494 (AF005048) transport inhibitor

response 1 [Arabidopsis thaliana]

Seq. No. 6609

Contig ID 7070 1.R1040 5'-most EST kmv700741748.h1

Method BLASTX
NCBI GI g399940
BLAST score 1968
E value 0.0e+00
Match length 451
% identity 96

NCBI Description MITOCHONDRIAL HEAT SHOCK 70 KD PROTEIN PRECURSOR

>gi_100004_pir__S25005 heat shock protein, 70K - kidney bean >gi_22636_emb_CAA47345_ (X66874) 70 kDa heat shock

protein [Phaseolus vulgaris]

Seq. No. 6610

Contig ID 7070 2.R1040 5'-most EST hrw701060186.h1

Method BLASTN NCBI GI g20834 BLAST score 204

```
E value
                  1.0e-111
Match length
                  400
% identity
                  88
NCBI Description
                  P.sativum PHSP1 mRNA for HSP70
                  6611
Seq. No.
                  7072 1.R1040
Contig ID
5'-most EST
                  LIB3170-034-Q1-K1-F4
Method
                  BLASTX
NCBI GI
                  g2462744
BLAST score
                  291
                  4.0e-26
E value
Match length
                  86
                  65
% identity
                  (AC002292) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  6612
Contig ID
                  7073 1.R1040
                  pcp700994935.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1304599
BLAST score
                  307
E value
                  1.0e-27
Match length
                  222
% identity
                  26
NCBI Description (U41315) ZNF127-Xp [Homo sapiens]
Seq. No.
                  6613
                  7073 2.R1040
Contig ID
                  fde700872991.hl
5'-most EST
                  6614
Seq. No.
Contig ID
                  7079 1.R1040
5'-most EST
                  LIB3170-032-Q1-K1-F3
Method
                  BLASTX
NCBI GI
                  g3021336
BLAST score
                  261
E value
                  2.0e-22
Match length
                  122
% identity
                  50
NCBI Description (AJ224957) RGA-like [Arabidopsis thaliana]
Seq. No.
                  6615
```

Contig ID

7079 2.R1040

5'-most EST

LIB3039-051-Q1-E1-F6

Method NCBI GI BLAST score BLASTX g3785986

E value

265 6.0e-23

Match length

111

% identity

50

NCBI Description

(AC005560) RGA1 protein [Arabidopsis thaliana]

Seq. No.

6616

Contig ID

7080 1.R1040

5'-most EST

LIB3030-005-Q1-B1-H8



Seq. No. 6617

Contig ID 7086_1.R1040 5'-most EST jsh701069332.h1

Seq. No. 6618

Contig ID 7091 1.R1040

5'-most EST LIB3107-032-Q1-K1-C5

Method BLASTN
NCBI GI g4415928
BLAST score 51
E value 2.0e-19
Match length 103
% identity 87

NCBI Description Arabidopsis thaliana chromosome II BAC F13A10 genomic sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 6619

Contig ID 7091_2.R1040

5'-most EST uC-gmropic098e09b1

Method BLASTX
NCBI GI g4415936
BLAST score 255
E value 8.0e-22
Match length 184
% identity 36

NCBI Description (AC006418) unknown protein [Arabidopsis thaliana]

Seq. No. 6620

Contig ID 7099 1.R1040

5'-most EST LIB3039-009-Q1-E1-G3

Method BLASTX
NCBI GI g3281861
BLAST score 194
E value 1.0e-14
Match length 122
% identity 39

NCBI Description (AL031004) putative protein [Arabidopsis thaliana]

Seq. No. 6621

Contig ID 7100 1.R1040

5'-most EST jC-gmle01810056d04a1

Method BLASTX
NCBI GI g4115379
BLAST score 309
E value 3.0e-28
Match length 101
% identity 59

NCBI Description (AC005967) putative carbonyl reductase [Arabidopsis

thaliana]

Seq. No. 6622

Contig ID 7100 2.R1040 5'-most EST zsg701119727.h1

Method BLASTX
NCBI GI g4115379
BLAST score 287
E value 1.0e-25



```
Match length
% identity
                   58
NCBI Description
                   (AC005967) putative carbonyl reductase [Arabidopsis
                   thaliana]
Seq. No.
                   6623
                   7100 3.R1040
Contig ID
5'-most EST
                   LIB3030-005-Q1-B1-G6
Method
                   BLASTX
NCBI GI
                   g4115379
BLAST score
                   239
E value
                   4.0e-20
Match length
                   84
                   56
% identity
NCBI Description
                   (AC005967) putative carbonyl reductase [Arabidopsis
                   thaliana]
Seq. No.
                   6624
                   7100 5.R1040
Contig ID
                   leu7\overline{0}1145195.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4115379
BLAST score
                   174
E value
                   1.0e-12
Match length
                   51
% identity
                   65
                   (AC005967) putative carbonyl reductase [Arabidopsis
NCBI Description
                   thaliana]
                   6625
Seq. No.
                   7103 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220073c07a1
Method
                   BLASTX
NCBI GI
                   g1421730
BLAST score
                   281
E value
                   7.0e-25
Match length
                   92
% identity
                   58
NCBI Description (U43082) RF2 [Zea mays]
                   6626
Seq. No.
                   7103 2.R1040
Contig ID
5'-most EST
                   uC-gmropic062f11b1
                   BLASTX
Method
NCBI GI
                   g1421730
BLAST score
                   167
                   2.0e-11
E value
Match length
                   75
% identity
NCBI Description (U43082) RF2 [Zea mays]
Seq. No.
                   6627
```

Contig ID 7107_1.R1040 5'-most EST leu701155568.h1 Method BLASTX

NCBI GI g3559807 BLAST score 588

E value 4.0e-61 Match length 124 % identity 60

NCBI Description (Y15628) HCF136 protein [Arabidopsis thaliana]

Seq. No. 6628

Contig ID 7108_1.R1040 5'-most EST vzy700752446.h1

Seq. No. 6629

Contig ID 7108 2.R1040

5'-most EST jC-gmle01810026a04a1

Seq. No. 6630

Contig ID 7112 1.R1040

5'-most EST LIB3030-005-Q1-B1-E7

Seq. No. 6631

Contig ID 7114 1.R1040 5'-most EST fua701039921.h1

Method BLASTN
NCBI GI 9438130
BLAST score 237
E value 1.0e-130
Match length 673
% identity 84

NCBI Description S.tuberosum mitochondrial gene for the 59kDa subunit of the

NAD+-dependent malic enzyme

Seq. No. 6632

Contig ID 7114_2.R1040

5'-most EST LIB3030-005-Q1-B1-E9

Method BLASTX
NCBI GI g585452
BLAST score 715
E value 1.0e-75
Match length 168
% identity 80

NCBI Description MALATE OXIDOREDUCTASE (NAD), MITOCHONDRIAL 59 KD ISOFORM

PRECURSOR (MALIC ENZYME) (ME) (NAD-DEPENDENT MALIC ENZYME) (NAD-ME) >gi_1076666_pir_ A53318 malate dehydrogenase (decarboxylating) (EC 1.1.1.39) 59K chain precursor, mitochondrial - potato >gi_438131_emb_CAA80547_ (Z23002) precursor of the 59kDa subunit of the mitochondrial

NAD+-dependent malic enzyme [Solanum tuberosum]

Seq. No. 6633

Contig ID 7117 1.R1040

5'-most EST LIB3030-005-Q1-B1-F2

Seq. No. 6634

Contig ID 7121_1.R1040 5'-most EST jex700907344.h1

Seq. No. 6635

Contig ID 7123 1.R1040 5'-most EST fua701042593.h1



Method BLASTX
NCBI GI g1703036
BLAST score 189
E value 5.0e-14
Match length 79
% identity 47

NCBI Description 5'-AMP-ACTIVATED PROTEIN KINASE, BETA-1 SUBUNIT (AMPK BETA-1 CHAIN) (40 KD SUBUNIT) >qi 1335858 (U42411)

5'-AMP-activated protein kinase, beta subunit [Rattus

norvegicus]

Seq. No. 6636

Contig ID 7123 2.R1040

5'-most EST uC-gmflminsoy026h11b1

Seq. No. 6637

Contig ID 7124 1.R1040 5'-most EST leu701150036.h1

Method BLASTN
NCBI GI g18569
BLAST score 425
E value 0.0e+00
Match length 476
% identity 100

NCBI Description G.max coxII gene for cytochrome oxidase subunit

Seq. No. 6638

Contig ID 7125 1.R1040

5'-most EST LIB3030-005-Q1-B1-C9

Method BLASTX
NCBI GI g4263777
BLAST score 1389
E value 1.0e-154
Match length 348
% identity 70

NCBI Description (AC006068) putative serine carboxypeptidase II [Arabidopsis

thaliana] >gi_4510391_gb_AAD21479.1 (AC007017) putative

serine carboxypeptidase II [Arabidopsis thaliana]

Seq. No. 6639

Contig ID 7127 1.R1040

5'-most EST uC-gmropic037g05b1

Method BLASTX
NCBI GI g2344887
BLAST score 740
E value 2.0e-78
Match length 223
% identity 66

NCBI Description (AC002388) unknown protein [Arabidopsis thaliana]

Seq. No. 6640

Contig ID 7127 2.R1040

5'-most EST LIB3030-005-Q1-B1-D10

Seq. No. 6641

Contig ID 7129 1.R1040 5'-most EST uaw700665131.h1



```
6642
Seq. No.
Contig ID
                   7129 2.R1040
5'-most EST
                   zhf700962637.h1
Seq. No.
                   6643
                   7132 1.R1040
Contig ID
5'-most EST
                   LIB3051-106-Q1-K1-H12
Method
                   BLASTN
NCBI GI
                   q975703
BLAST score
                   42
                   3.0e-14
E value
Match length
                   85
% identity
                   54
NCBI Description P.sativum GR gene
Seq. No.
                   6644
Contig ID
                   7132 2.R1040
5'-most EST
                   q560\overline{5}768
Method
                   BLASTX
NCBI GI
                   q2262156
BLAST score
                   163
                   8.0e-12
E value
                   115
Match length
% identity
                   43
NCBI Description
                   (AC002329) hypothetical protein with weak similarity to S.
                   cerevisiae hypothetical protein PID:g1066493 [Arabidopsis
                   thaliana]
Seq. No.
                   6645
Contig ID
                   7135 2.R1040
5'-most EST
                   sat701003967.h1
Method
                   BLASTX
                   g4204695
NCBI GI
BLAST score
                   153
                   5.0e-10
E value
Match length
                   98
                   38
% identity
NCBI Description
                   (AF117062) putative inositol polyphosphate 5-phosphatase
                   At5P1 [Arabidopsis thaliana]
Seq. No.
                   6646
Contig ID
                   7136 1.R1040
5'-most EST
                   LIB3030-005-Q1-B1-D9
Method
                   BLASTX
NCBI GI
                   g2642153
BLAST score
                   352
                   7.0e-33
E value
Match length
                   206
% identity
                   39
```

NCBI Description

(AC003000) unknown protein [Arabidopsis thaliana] >gi 2795810 (AC003674) unknown protein [Arabidopsis

thaliana]

Seq. No.

6647

Contig ID 7136_2.R1040

5'-most EST

jC-gmro02910044e04a1



Seq. No. 6648 Contig ID 7138

Contig ID 7138_1.R1040
5'-most EST g5606078
Method BLASTX
NCBI GI g3023751
BLAST score 440
E value 2.0e-43
Match length 100
% identity 28

NCBI Description 70 KD PEPTIDYLPROLYL ISOMERASE (PEPTIDYLPROLYL CIS-TRANS ISOMERASE) (CYCLOPHILIN) (PPIASE) >gi_1076772_pir S55383

peptidylprolyl isomerase (EC 5.2.1.8) - wheat

>gi_854626_emb_CAA60505_ (X86903) peptidylprolyl isomerase

[Triticum aestivum]

Seq. No. 6649

Contig ID 7139 1.R1040

5'-most EST LIB3030-005-Q1-B1-E11

Method BLASTX
NCBI GI g4371285
BLAST score 398
E value 9.0e-39
Match length 104
% identity 71

NCBI Description (AC006260) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6650

Contig ID 7140 1.R1040 5'-most EST g4282509 Method BLASTX NCBI GI g3702368 BLAST score 1123 E value 1.0e-123 Match length 296 % identity 68

NCBI Description (AJ001855) alpha subunit of F-actin capping protein

[Arabidopsis thaliana]

Seq. No. 6651

Contig ID 7141_1.R1040 5'-most EST epx701106727.h1

Seq. No. 6652

Contig ID 7146 1.R1040

5'-most EST uC-gmronoir064d10b1

Seq. No. 6653

Contig ID 7147_1.R1040 5'-most EST fde700874693.h1

Method BLASTX
NCBI GI g3080437
BLAST score 180
E value 6.0e-13
Match length 54
% identity 61

NCBI Description (AL022605) putative protein [Arabidopsis thaliana]



Seq. No. 6654 Contig ID 7152 1.R1040 5'-most EST jC-gmst02400009e02a1 Method BLASTX NCBI GI g3269288 BLAST score 800 E value 2.0e-85 Match length 199 % identity 78 NCBI Description (AL030978) putative protein [Arabidopsis thaliana] Seq. No. 6655 Contig ID 7153 1.R1040 5'-most EST rca700996111.hl Method BLASTX NCBI GI g3450842 BLAST score 624 E value 4.0e-65 Match length 159 % identity 68 (AF080436) mitogen activated protein kinase kinase [Oryza NCBI Description sativa] Seq. No. 6656 Contig ID 7154 1.R1040 5'-most EST LIB3030-005-Q1-B1-C4 Method BLASTX NCBI GI q4455192 BLAST score 376 E value 4.0e-36 Match length 150 % identity 49 NCBI Description (AL035440) putative protein [Arabidopsis thaliana] Seq. No. 6657 Contig ID 7156 1.R1040 5'-most EST xpa700792358.h1Method BLASTX NCBI GI g3142295 BLAST score 249

E value 4.0e-21 Match length 94 % identity 53

NCBI Description (AC002411) Strong similarity to phosphoribosylanthranilate

transferase gb_D86180 from Pisum sativum. [Arabidopsis

thaliana]

Seq. No. 6658

Contig ID 7156 2.R1040 5'-most EST vzy700754349.h1

Method BLASTX NCBI GI q3142295 BLAST score 176 E value 1.0e-12 Match length 53 % identity 62



NCBI Description (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb_D86180 from Pisum sativum. [Arabidopsis thaliana]

Seq. No. 6659

Contig ID 7160_1.R1040 5'-most EST jex700905076.h1

Method BLASTX
NCBI GI g2924512
BLAST score 355
E value 1.0e-33
Match length 91
% identity 67

NCBI Description (AL022023) beta-galactosidase - like protein [Arabidopsis

thaliana]

Seq. No. 6660

Contig ID 7164 1.R1040

5'-most EST uC-gmrominsoy0001h07b1

Seq. No. 6661

Contig ID 7167 1.R1040

5'-most EST LIB3030-005-Q1-B1-A2

Seq. No. 6662

Contig ID 7170 1.R1040

5'-most EST LIB3030-005-Q1-B1-A6

Seq. No. 6663

Contig ID 7174_1.R1040 5'-most EST zhf700962284.h1

Method BLASTX
NCBI GI g4467128
BLAST score 897
E value 6.0e-97
Match length 228
% identity 76

NCBI Description (AL035538) putative protein [Arabidopsis thaliana]

Seq. No. 6664

Contig ID 7176_1.R1040

5'-most EST LIB3030-004-Q1-B1-G1

Method BLASTX
NCBI GI g4512652
BLAST score 169
E value 1.0e-11
Match length 221
% identity 28

NCBI Description (AC007048) hypothetical protein [Arabidopsis thaliana]

Seq. No. 6665

Contig ID 7177 1.R1040

5'-most EST uC-gmflminsoy080c10b1

Method BLASTX
NCBI GI g2213590
BLAST score 421
E value 4.0e-41



Match length 158 % identity 51

NCBI Description (AC000348) T7N9.10 [Arabidopsis thaliana]

Seq. No.

6666

Contig ID

7179 1.R1040

5'-most EST

LIB3030-004-Q1-B1-G4

Method BLASTX NCBI GI g4455359 BLAST score 870 E value 1.0e-93 Match length 261 % identity 70

NCBI Description (AL035524) putative protein [Arabidopsis thaliana]

Seq. No.

6667

Contig ID 5'-most EST 7182 1.R1040 zsq701123319.h1

Seq. No.

6668

Contig ID 5'-most EST 7186 1.R1040 epx701107104.h1

Method BLASTX g2213590 NCBI GI BLAST score 178 E value 6.0e-13 Match length 86

% identity 41

NCBI Description (AC000348) T7N9.10 [Arabidopsis thaliana]

Seq. No.

6669

Contig ID

7192 1.R1040

5'-most EST

jC-gmle01810080f03a1

Seq. No.

6670

Contig ID

7193 1.R1040

5'-most EST

LIB3109-042-Q1-K1-H1

Method BLASTX g1346766 NCBI GI BLAST score 369 9.0e-35 E value Match length 257 % identity 35

NCBI Description

26S PROTEASOME REGULATORY SUBUNIT S14 (P31)

>gi_1362741_pir__S56108 multicatalytic endopeptidase complex (EC 3.4.99.46) regulatory chain 31 - human

>gi_1037164_dbj_BAA07237 (D38047) 26S proteasome subunit p31 [Homo sapiens] >gi_3702282 (AC005789) PP31 HUMAN [Homo sapiens] >gi_4506233_ref_NP_002803.1_pPSMD8_ proteasome

(prosome, macropain) 26S subunit, non-ATPase,

Seq. No.

6671

Contig ID

7194 1.R1040

5'-most EST

LIB3030-004-Q1-B1-E5

Method BLASTX NCBI GI q3068705 BLAST score 228



E value 7.0e-19 Match length 105 % identity 48

NCBI Description (AF049236) unknown [Arabidopsis thaliana]

Seq. No.

6672

Contig ID

7196 1.R1040

5'-most EST

LIB3030-004-Q1-B1-E8

Seq. No.

6673

Contig ID 5'-most EST

7197_1.R1040 leu701149726.h1

Seq. No.

6674

Contig ID

7203 1.R1040

5'-most EST

LIB3040-023-Q1-E1-A12

Method NCBI GI BLAST score BLASTN g1272409 240

E value
Match length
% identity

1.0e-132 384

NCBI Description

Vicia faba immunophilin precursor (FKBP15) mRNA, complete

cds

91

Seq. No.

6675

Contig ID

7209_1.R1040

5'-most EST

LIB3106-105-Q1-K1-E4

Method BLASTX
NCBI GI g2275202
BLAST score 178
E value 6.0e-13
Match length 60
% identity 60

NCBI Description

(AC002337) acyl-CoA synthetase isolog [Arabidopsis

thaliana]

Seq. No.

6676

Contig ID 5'-most EST

7213_1.R1040 bth700843983.h1

Method BLASTX
NCBI GI g1841475
BLAST score 353
E value 2.0e-33
Match length 80
% identity 75

NCBI Description (Y11105) Myb26 [Pisum sativum]

Seq. No.

6677

Contig ID

7213 2.R1040

5'-most EST

uC-gmflminsoy037e05b1

Method BLASTX
NCBI GI g1002796
BLAST score 500
E value 9.0e-51
Match length 121
% identity 74

1244

% identity

58



```
(U33915) Cpm10 [Craterostigma plantagineum]
NCBI Description
                   6678
Seq. No.
Contig ID
                   7214 1.R1040
5'-most EST
                   rrt700646076.h1
Method
                  BLASTX
NCBI GI
                   q2130139
BLAST score
                   162
                   5.0e-11
E value
Match length
                   99
                   39
% identity
NCBI Description
                  hypothetical protein - maize >gi 949980 emb CAA61258
                   (X88779) orf [Zea mays]
Seq. No.
                   6679
Contig ID
                  7215 1.R1040
5'-most EST
                   zhf700960909.h1
Method
                  BLASTX
NCBI GI
                  q2765366
BLAST score
                  231
E value
                  3.0e-19
Match length
                  99
% identity
                  45
NCBI Description
                  (Y14038) putative Ole e 1 protein [Betula pendula]
Seq. No.
                  6680
Contig ID
                  7217 1.R1040
                  LIB3087-002-Q1-K1-D3
5'-most EST
Method
                  BLASTX
                  g2880048
NCBI GI
                  310
BLAST score
E value
                  5.0e-28
Match length
                  71
% identity
                  79
NCBI Description
                  (AC002340) unknown protein [Arabidopsis thaliana]
Seq. No.
                  6681
Contig ID
                  7218 1.R1040
5'-most EST
                  uC-gmflminsoy065f10b1
Method
                  BLASTX
NCBI GI
                  q3152605
BLAST score
                  277
E value
                  2.0e-24
Match length
                  134
% identity
                  49
NCBI Description
                  (AC004482) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  6682
                  7221 1.R1040
Contig ID
5'-most EST
                  V4R-02-Q1-B1-C6
Method
                  BLASTX
NCBI GI
                  q4539335
BLAST score
                  1046
E value
                  1.0e-114
Match length
                  315
```

NCBI Description (AL035539) putative protein [Arabidopsis thaliana]



```
Seq. No.
                  6683
Contig ID
                  7222 1.R1040
5'-most EST
                  V4R-01-Q1-B1-G5
Method
                  BLASTX
NCBI GI
                  g1168470
BLAST score
                  1543
E value
                  1.0e-172
Match length
                  421
% identity
                  73
NCBI Description
                  PROTEIN KINASE APK1A >qi 282877 pir S28615 protein kinase,
                  tyrosine/serine/threonine-specific (EC 2.7.1.-) -
                  Arabidopsis thaliana >gi 217829 dbj BAA02092 (D12522)
                  protein tyrosine-serine-threonine kinase [Arabidopsis
                  thaliana]
Seq. No.
                  6684
Contig ID
                  7222 2.R1040
5'-most EST
                  zsq701117356.h1
Method
                  BLASTX
NCBI GI
                  q3461835
BLAST score
                  343
E value
                  5.0e-32
Match length
                  194
% identity
NCBI Description
                  (AC005315) putative protein kinase [Arabidopsis thaliana]
                  >gi_3927840 (AC005727) putative protein kinase [Arabidopsis
                  thaliana]
Seq. No.
                  6685
Contig ID
                  7222 5.R1040
5'-most EST
                  eep700865844.h1
Seq. No.
                  6686
Contig ID
                  7226 1.R1040
5'-most EST
                  uC-gmflminsoy028g12b1
Method
                  BLASTX
                  g2493046
NCBI GI
BLAST score
                  735
E value
                  9.0e-78
Match length
                  197
                  75
% identity
                  ATP SYNTHASE DELTA' CHAIN, MITOCHONDRIAL PRECURSOR
NCBI Description
                  >gi 82297 pir A41740 H+-transporting ATP synthase (EC
                  3.6.1.34) delta' chain precursor - sweet potato
                  >gi_217938_dbj_BAA01511_ (D10660) mitochondrial F1-ATPase
                  delta subunit [Ipomoea batatas]
Seq. No.
                  6687
```

 Seq. No.
 6687

 Contig ID
 7229_1.R1040

 5'-most EST
 leu701156154.h1

 Method
 BLASTX

 NCBI GI
 g1848212

 BLAST score
 751

 E value
 7.0e-80

 Match length
 181

47

% identity



```
NCBI Description (Y11209) protein disulfide-isomerase precursor [Nicotiana tabacum]
```

Seq. No. 6688

Contig ID 7230_1.R1040

5'-most EST jC-gmst02400076d08a1

Method BLASTN
NCBI GI g1556445
BLAST score 224
E value 1.0e-122
Match length 608
% identity 84

NCBI Description Hordeum vulgare alpha tubulin (tubA) mRNA, complete cds

Seq. No. 6689

Contig ID 7231_1.R1040 5'-most EST asn701132470.h1

Method BLASTX
NCBI GI g2388580
BLAST score 1259
E value 1.0e-139
Match length 252
% identity 90

NCBI Description (AC000098) Similar to Sequence 10 from patent 5477002

(gb_1253956). [Arabidopsis thaliana]

Seq. No. 6690

Contig ID 7233 1.R1040

5'-most EST LIB3030-004-Q1-B1-C1

Method BLASTX
NCBI GI g542050
BLAST score 325
E value 5.0e-30
Match length 140
% identity 46

NCBI Description catechol O-methyltransferase (EC 2.1.1.6) III - common

tobacco >gi 429114 emb CAA50561 (X71430) catechol

O-methyltransferase [Nicotiana tabacum]

Seq. No. 6691

Contig ID 7235_1.R1040 5'-most EST gsv701045787.h1

Seq. No. 6692

Contig ID 7239 1.R1040

5'-most EST LIB3030-003-Q1-B1-G9

Seq. No. 6693

Contig ID 7240 1.R1040

5'-most EST LIB3 $\overline{1}$ 07-031-Q1-K1-F5

Seq. No. 6694

Contig ID 7240_2.R1040 5'-most EST sat701007364.h2

Seq. No. 6695

Contig ID 7241_1.R1040

% identity

82



```
ssr700559125.h1
5'-most EST
Method
                  BLASTN
                  g887572
NCBI GI
BLAST score
                  74
                  3.0e-33
E value
                  206
Match length
                  84
% identity
NCBI Description P.sativum mRNA for starch synthase (2988 bp)
Seq. No.
                  6696
                  7245 1.R1040
Contig ID
5'-most EST
                  bth700846389.h1
                                                               130cm ^
Method
                  BLASTX
                  q3023713
NCBI GI
BLAST score
                  642
E value
                  3.0e-67
Match length
                  141
% identity
                  87
                  ENOLASE (2-PHOSPHOGLYCERATE DEHYDRATASE)
NCBI Description
                  (2-PHOSPHO-D-GLYCERATE HYDRO-LYASE) (OSE1) >gi 780372
                  (U09450) enolase [Oryza sativa]
                  6697
Seq. No.
Contig ID
                  7245 2.R1040
5'-most EST
                  uC-gmropic045a01b1
Method
                  BLASTX
NCBI GI
                  q3023713
BLAST score
                  478
E value
                  4.0e-48
Match length
                  110
% identity
                  85
                  ENOLASE (2-PHOSPHOGLYCERATE DEHYDRATASE)
NCBI Description
                  (2-PHOSPHO-D-GLYCERATE HYDRO-LYASE) (OSE1) >gi 780372
                  (U09450) enolase [Oryza sativa]
                  6698
Seq. No.
Contig ID
                  7247 1.R1040
                  LIB3170-043-Q1-J1-G7
5'-most EST
Method
                  BLASTN
NCBI GI
                  q4159706
BLAST score
                  49
                  2.0e-18
E value
Match length
                  173
% identity
NCBI Description
                  Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:
                  MGL6, complete sequence [Arabidopsis thaliana]
                  6699
Seq. No.
                  7247 2.R1040
Contig ID
5'-most EST
                  leu701149612.h1
Method
                  BLASTN
NCBI GI
                  q4159706
BLAST score
                  52
                  5.0e-20
E value
Match length
                  176
```

1248

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:



MGL6, complete sequence [Arabidopsis thaliana]

Seq. No. 6700

Contig ID 7247 3.R1040

5'-most EST LIB3106-110-Q1-K1-C9

83

Method BLASTN NCBI GI q4159706 BLAST score 41 2.0e-13 E value 125 Match length

% identity Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: NCBI Description

MGL6, complete sequence [Arabidopsis thaliana]

6701 Seq. No.

Contig ID 7247 4.R1040 5'-most EST dpv701097991.h1

Method BLASTN NCBI GI q4159706 BLAST score 35 5.0e-10 E value Match length 131 % identity

Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: NCBI Description

MGL6, complete sequence [Arabidopsis thaliana]

Seq. No. 6702

Contig ID 7247 5.R1040

5'-most EST LIB3138-127-Q1-N1-D11

Seq. No. 6703

Contig ID 7250 1.R1040

5'-most EST LIB3040-002-Q1-E1-B3

Method BLASTX g3721856 NCBI GI BLAST score 208 2.0e-16 E value Match length 54

65 % identity

NCBI Description (AB014057) beta-Amyrin Synthase [Panax ginseng]

Seq. No. 6704

7251 1.R1040 Contig ID

5'-most EST uC-gmropic091b10b1

6705 Seq. No.

7253 1.R1040 Contig ID

5'-most EST uC-gmflminsoy068e02b1

6706 Seq. No.

7259 1.R1040 Contig ID

5'-most EST uC-gmflminsoy044b04b1

Method BLASTX q4049341 NCBI GI BLAST score 585 E value 1.0e-60 Match length 140

Seq. No.

Contig ID

6712

7267 1.R1040



```
% identity
                  (AL034567) putative protein [Arabidopsis thaliana]
NCBI Description
                   6707
Seq. No.
                  7260 1.R1040
Contig ID
5'-most EST
                  LIB3030-003-Q1-B1-F12
Method
                  BLASTX
NCBI GI
                  q99733
BLAST score
                  806
                  1.0e-96
E value
                  256
Match length
% identity
                   72
                  hypothetical protein Hlflk - Arabidopsis thaliana
NCBI Description
                   (fragment) >gi 16482_emb_CAA44318_ (X62461) H1flk
                   [Arabidopsis thaliana]
Seq. No.
                   6708
                   7262 1.R1040
Contig ID
                  sat701004636.h1
5'-most EST
Method
                  BLASTX
                  q2894607
NCBI GI
BLAST score
                   436
                   5.0e-43
E value
                   150
Match length
% identity
                   54
                  (AL021889) NAM (no apical meristem) - like protein
NCBI Description
                   [Arabidopsis thaliana]
                   6709
Seq. No.
Contig ID
                   7262 2.R1040
5'-most EST
                   LIB3051-059-Q1-K2-D5
Method
                   BLASTX
NCBI GI
                   q2894607
BLAST score
                   199
                - 2.0e-15
E value
Match length
                   109
% identity
                   39
                  (AL021889) NAM (no apical meristem) - like protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   6710
Contig ID
                   7264 1.R1040
5'-most EST
                   LIB3107-041-Q1-K1-B5
Method
                   BLASTX
                   g3747111
NCBI GI
BLAST score
                   661
                   6.0e-69
E value
Match length
                   226
% identity
                   56
NCBI Description
                  (AF095641) MTN3 homolog [Arabidopsis thaliana]
Seq. No.
                   6711
                   7264 2.R1040
Contig ID
                   LIB3<u>1</u>70-082-Q1-J1-F1
5'-most EST
```



```
zzp700830203.h1
5'-most EST
Seq. No.
                  6713
Contig ID
                  7269 1.R1040
5'-most EST
                  jC-gmro02910059c03d1
Seq. No.
                  6714
Contig ID
                  7269 2.R1040
5'-most EST
                  uC-gmrominsoy026g03b1
Seq. No.
                  6715
Contig ID
                  7271 1.R1040
5'-most EST
                  LIB3138-125-Q1-N1-B7
Seq. No.
                  6716
Contig ID
                  7271 2.R1040
5'-most EST
                  LIB3030-003-Q1-B1-C12
Method
                  BLASTX
NCBI GI
                  q1742187
BLAST score
                  185
E value
                  3.0e-13
Match length
                  150
% identity
                  30
                  (D90771) ORF ID:o260#14; similar to [SwissProt Accession
NCBI Description
                  Number P11666 [Escherichia coli] >gi 1742198 dbj BAA14933
                   (D90772) ORF ID:o260#14; similar to [SwissProt Accession
                  Number P1166\overline{6} [Escherichia coli] >gi 1787591 (AE000231)
                  orf, hypothetical protein [Escherichia coli]
                  6717
Seq. No.
                  7276 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy177g01b1
Method
                  BLASTX
NCBI GI
                  g2832683
BLAST score
                  673
E value
                  1.0e-70
Match length
                  160
% identity
                  74
NCBI Description (AL021712) putative protein [Arabidopsis thaliana]
Seq. No.
                   6718
                  7277 1.R1040
Contig ID
5'-most EST
                   jC-gmst02400055b10a1
Method
                  BLASTX
NCBI GI
                  g3298443
BLAST score
                   641
```

E value

Match length

5.0e-67

% identity

184 73

NCBI Description

(AB010880) chloroplast ribosomal protein L17 [Nicotiana

tabacum]

Seq. No.

6719

6720

Contig ID

7277 2.R1040

5'-most EST

LIB3030-003-Q1-B1-D11

Seq. No.



7277 3.R1040 Contig ID

LIB3106-072-P1-K1-G1 5'-most EST

Method BLASTX NCBI GI g3298443 BLAST score 155 E value 2.0e-10 Match length 57 % identity 61

(AB010880) chloroplast ribosomal protein L17 [Nicotiana NCBI Description

tabacum]

Seq. No. 6721

7278 2.R1040 Contig ID

5'-most EST LIB3030-003-Q1-B1-D12

Seq. No. 6722

Contig ID 7280 1.R1040

5'-most EST LIB3030-003-Q1-B1-D5

Method BLASTX NCBI GI q4206122 BLAST score 419 E value 3.0e-41Match length 146 60 % identity

(AF097667) protein phosphatase 2C homolog [Mesembryanthemum NCBI Description

crystallinum]

Seq. No. 6723

7281 1.R1040 Contig ID

5'-most EST uC-gmropic050b08b1

BLASTX Method NCBI GI g399333 BLAST score 489 2.0e-49 E value Match length 128 % identity

NCBI Description CYSTEINE SYNTHASE CHLOROPLAST PRECURSOR (O-ACETYLSERINE SULFHYDRYLASE) (O-ACETYLSERINE (THIOL)-LYASE) (CSASE)

>gi 322740 pir A43407 cysteine synthase (EC 4.2.99.8) precursor - pepper >gi 17944 emb CAA46086 (X64874)

O-acetylserine (thiol)-lyase [Capsicum annuum]

Seq. No. 6724

7281 2.R1040 Contig ID

LIB3030-003-Q1-B1-D7 5'-most EST

Seq. No. 6725

Contig ID 7286 1.R1040

jC-gmle01810055g05a1 5'-most EST

Method BLASTX g4006865 NCBI GI BLAST score 817 2.0e-87 E value Match length 227 % identity 67

(Z99707) putative protein [Arabidopsis thaliana] NCBI Description



```
6726
Seq. No.
Contig ID
                   7286 3.R1040
5'-most EST
                   zpv700761628.h1
                   BLASTX
Method
NCBI GI
                   g2190551
BLAST score
                   163
E value
                   2.0e-11
                   39
Match length
% identity
                   (AC001229) Similar to C. elegans hypothetical protein
NCBI Description
                   K07C5.6 (gb_Z71181). ESTs gb H36844,gb AA394956 come from
                   this gene. [Arabidopsis thaliana]
                  6727
Seq. No.
                   7287 1.R1040
Contig ID
5'-most EST
                   leu7\overline{0}1147472.h1
Method
                   BLASTX
NCBI GI
                   q541944
BLAST score
                   632
E value
                   1.0e-65
Match length
                   181
% identity
                   71
NCBI Description
                  nodulin-26 - soybean
                   6728
Seq. No.
Contig ID
                   7287 2.R1040
5'-most EST
                   asn701134248.h1
Method
                   BLASTX
NCBI GI
                   g1076748
                   379
BLAST score
                   8.0e-37
E value
Match length
                   93
                   77
% identity
                   major intrinsic protein - rice >gi_440869_dbj_BAA04257_
NCBI Description
                   (D17443) major intrinsic protein [Oryza sativa]
                   6729
Seq. No.
                   7289 1.R1040
Contig ID
5'-most EST
                   bnh700764511.h1
                   BLASTX
Method
                   g1076668
NCBI GI
BLAST score
                   650
E value
                   4.0e-68
Match length
                   158
                   79
% identity
                   NADH dehydrogenase (EC 1.6.99.3) - potato
NCBI Description
                   >gi 639834 emb CAA58823 (X83999) NADH dehydrogenase
                   [Solanum tuberosum]
```

Seq. No. 6730

Contig ID 7289 2.R1040 5'-most EST leu701151520.h1

Method BLASTX
NCBI GI g1076668
BLAST score 153
E value 2.0e-10
Match length 62



% identity 60

NCBI Description NADH dehydrogenase (EC 1.6.99.3) - potato

>gi 639834 emb CAA58823_ (X83999) NADH dehydrogenase

[Solanum tuberosum]

Seq. No. 6731

Contig ID 7290_1.R1040 5'-most EST dpv701097134.h1

Seq. No. 6732

Contig ID 7290 2.R1040

5'-most EST LIB3093-015-Q1-K1-C8

Seq. No. 6733

Contig ID 7291_1.R1040 5'-most EST kl1701207653.h1

Seq. No. 6734

Contig ID 7291_2.R1040 5'-most EST pmv700894835.h1

Seq. No. 6735

Contig ID 7294_1.R1040 5'-most EST jex700906742.h1

Method BLASTX
NCBI GI g2529703
BLAST score 300
E value 6.0e-27
Match length 159
% identity 53

NCBI Description (AF000142) class II knotted-like homeodomain protein

[Lycopersicon esculentum]

Seq. No. 6736

Contig ID 7294_2.R1040 5'-most EST sat701003783.h1

Method BLASTX
NCBI GI g1805618
BLAST score 190
E value 4.0e-14
Match length 126
% identity 47

NCBI Description (D49704) OSH45 transcript [Oryza sativa]

Seq. No. 6737

Contig ID 7295_1.R1040 5'-most EST wvk700684662.h1

Method BLASTX
NCBI GI g2462781
BLAST score 748
E value 2.0e-79
Match length 219
% identity 69

NCBI Description (U73175) carbamoyl phosphate synthetase small subunit

[Arabidopsis thaliana]

Seq. No. 6738

1254



Contig ID 7295_2.R1040 5'-most EST vwf700675431.h1

Method BLASTX
NCBI GI g2462781
BLAST score 522
E value 4.0e-53
Match length 142
% identity 75

NCBI Description (U73175) carbamoyl phosphate synthetase small subunit

[Arabidopsis thaliana]

Seq. No. 6739

Contig ID 7296_1.R1040 5'-most EST txt700734315.h1

Method BLASTX
NCBI GI g465740
BLAST score 261
E value 7.0e-22
Match length 374
% identity 24

NCBI Description HYPOTHETICAL 54.9 KD PROTEIN C02F5.7 IN CHROMOSOME III

>gi_630513_pir__S44609 hypothetical protein C02F5.7 Caenorhabditis elegans >gi_289614 (L14745) homology with
glucose induced repressor, GRR1; putative [Caenorhabditis

elegans]

Seq. No. 6740

Contig ID 7296_2.R1040 5'-most EST fua701037212.h1

Seq. No. 6741

Contig ID 7297_1.R1040

5'-most EST LIB3030-003-Q1-B1-B10

Method BLASTX
NCBI GI g4538975
BLAST score 289
E value 7.0e-26
Match length 62
% identity 85

NCBI Description (ALO49487) putative protein [Arabidopsis thaliana]

Seq. No. 6742

Contig ID 7305 1.R1040

5'-most EST LIB3139-080-P1-N1-E7

Method BLASTX
NCBI GI g4432839
BLAST score 454
E value 3.0e-45
Match length 90
% identity 52

NCBI Description (AC006283) unknown protein [Arabidopsis thaliana]

Seq. No. 6743

Contig ID 7305_2.R1040

5'-most EST uC-gmrominsoy108b05b1

Method BLASTN NCBI GI g4432829



BLAST score 38 E value 6.0e-12 Match length 94 % identity 85

NCBI Description Arabidopsis thaliana chromosome II BAC T1B3 genomic sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 6744

Contig ID 7308 1.R1040

5'-most EST LIB3049-037-Q1-E1-F1

Method BLASTX
NCBI GI 94235430
BLAST score 752
E value 1.0e-147
Match length 418
% identity 64

NCBI Description (AF098458) latex-abundant protein [Hevea brasiliensis]

Seq. No. 6745

Contig ID 7308_2.R1040 5'-most EST fde700875009.h1

Method BLASTX
NCBI GI g4235430
BLAST score 748
E value 2.0e-79
Match length 231
% identity 65

NCBI Description (AF098458) latex-abundant protein [Hevea brasiliensis]

Seq. No. 6746

Contig ID 7315 1.R1040

5'-most EST uC-gmrominsoy200b02b1

Method BLASTX
NCBI GI g2827529
BLAST score 793
E value 1.0e-84
Match length 304
% identity 53

NCBI Description (AL021633) putative protein [Arabidopsis thaliana]

Seq. No. 6747

Contig ID 7317_1.R1040

5'-most EST LIB3049-052-Q1-E1-H5

Seq. No. 6748

Contig ID 7319_1.R1040 5'-most EST hrw701057579.h1

Method BLASTX
NCBI GI g2244898
BLAST score 327
E value 4.0e-30
Match length 201
% identity 48

NCBI Description (Z97338) strong similarity to protein phosphatase 2A

regulatory chain, 74K [Arabidopsis thaliana]

Seq. No. 6749



```
Contig ID
                   7324 1.R1040
5'-most EST
                   LIB3106-040-Q1-K1-D10
Method
                   BLASTX
NCBI GI
                   g120663
BLAST score
                   1928
E value
                   0.0e + 00
Match length
                   442
% identity
                   83
NCBI Description
                   GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE B PRECURSOR,
                   CHLOROPLAST >gi 66027 pir DEPMNB
                   glyceraldehyde-3-phosphate dehydrogenase (NADP+)
                   (phosphorylating) (EC 1.2.1.13) B precursor, chloroplast -
                   garden pea >gi 309671 (M55147) glyceraldehyde-3-phosphate
                   dehydrogenase B subunit [Pisum sativum]
                   6750
Seq. No.
Contig ID
                   7324 2.R1040
5'-most EST
                   jC-gmle01810045f02d1
Method
                   BLASTN
NCBI GI
                   q20732
                   232
BLAST score
E value
                   1.0e-127
Match length
                   432
% identity
                   88
NCBI Description
                   Pea chloroplast GAPB mRNA encoding
                   glyceraldehyde-3-phosphate dehydrogenase (GAPDH) subunit B
                   (EC 1.2.1.13)
Seq. No.
                   6751
Contig ID
                   7324 3.R1040
5'-most EST
                   LIB3\overline{1}06-059-Q1-K1-E4
Method
                   BLASTN
NCBI GI
                   g20732
BLAST score
                   146
                   3.0e-76
E value
Match length
                   374
% identity
                   85
                   Pea chloroplast GAPB mRNA encoding
NCBI Description
                   glyceraldehyde-3-phosphate dehydrogenase (GAPDH) subunit B
                   (EC 1.2.1.13)
Seq. No.
                   6752
Contig ID
                   7324 5.R1040
5'-most EST
                   kmv700739659.h1
Method
                   BLASTN
NCBI GI
                   q20732
BLAST score
                   37
E value
                   2.0e-11
Match length
                   89
% identity
                   85
                  Pea chloroplast GAPB mRNA encoding
NCBI Description
                   glyceraldehyde-3-phosphate dehydrogenase (GAPDH) subunit B
```

(EC 1.2.1.13)

Seq. No.

6753 Contig ID 7324 6.R1040 5'-most EST zhf700961332.h1



Method BLASTX NCBI GI g2213610 BLAST score 146 E value 1.0e-09 Match length 47 % identity 32 NCBI Description (AC000103) F21J9.4 [Arabidopsis thaliana] Seq. No. 6754 Contig ID 7324 7.R1040 5'-most EST zhf700961396.h1 Method BLASTN NCBI GI g20732 BLAST score 43 4.0e-15 E value Match length 63 % identity 92 NCBI Description Pea chloroplast GAPB mRNA encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH) subunit B (EC 1.2.1.13) Seq. No. 6755 Contig ID 7326 1.R1040 5'-most EST LIB3039-015-Q1-E1-E2 Method BLASTX NCBI GI g4218120 BLAST score 739 E value 2.0e-99 Match length 260 % identity 72 NCBI Description (AL035353) Proline-rich APG-like protein [Arabidopsis thaliana] Seq. No. 6756 Contig ID 7327 1.R1040 5'-most EST LIB3092-007-Q1-K1-D7 Seq. No. 6757 Contig ID 7331 1.R1040 5'-most EST jex700904224.hl Method BLASTX NCBI GI g2505940 BLAST score 551 E value 1.0e-56 Match length 130 % identity NCBI Description (Y13071) 26S proteasome, non-ATPase subunit [Mus musculus] Seq. No. 6758 7331 2.R1040 Contig ID

5'-most EST LIB3087-012-Q1-K1-G3

Method BLASTX
NCBI GI g2505940
BLAST score 387
E value 2.0e-37
Match length 103
% identity 77



```
NCBI Description (Y13071) 26S proteasome, non-ATPase subunit [Mus musculus]
                  6759
Seq. No.
Contiq ID
                  7336 1.R1040
5'-most EST
                  zhf700956360.h1
Method
                  BLASTX
NCBI GI
                  q2828292
BLAST score
                  1556
E value
                  1.0e-174
Match length
                  404
% identity
                  71
NCBI Description
                  (AL021687) neoxanthin cleavage enzyme-like protein
                  [Arabidopsis thaliana]
Seq. No.
                  6760
Contig ID
                  7342 1.R1040
5'-most EST
                  LIB3056-013-Q1-N1-F6
Method
                  BLASTX
NCBI GI
                  g2500715
BLAST score
                  526
                  1.0e-53
E value
Match length
                  105
% identity
                  83
NCBI Description
                  PATHOGENESIS-RELATED PROTEIN PR-1 PRECURSOR
                  >gi 629627 pir S47171 gene PR-1 protein - barrel medic
                  >gi 505553 emb CAA56174 (X79778) PR-1 [Medicago
                  truncatula]
Seq. No.
                  6761
Contig ID
                  7343 1.R1040
5'-most EST
                  LIB3039-045-Q1-E1-D3
Seq. No.
                  6762
Contig ID
                  7348 1.R1040
5'-most EST
                  gsf700698427.h1
Seq. No.
                  6763
Contig ID
                  7351 1.R1040
5'-most EST
                  pxt700942441.h1
Method
                  BLASTX
NCBI GI
                  q3702102
BLAST score
                  259
E value
                  4.0e-30
Match length
                  134
% identity
                  51
                  (AL031765) 1-evidence=predicted by match;
NCBI Description
                  1-match accession=SPTREMBL:Q14105;
                  1-match_description=D1075-LIKE GENE (FRAGMENT).;
                  1-match species=HOMO SAPIENS (HUMAN).;
                  1-method=blastx; 1.4.9; 1-method score=101.00; 56.00;
                  180.00; 2
                  6764
Seq. No.
Contig ID
                  7355 1.R1040
```

5'-most EST uC-gmflminsoy032b07b1

Method BLASTN NCBI GI g4220637



BLAST score 54 5.0e-21 E value Match length 390 % identity 83 Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: NCBI Description MIE1, complete sequence [Arabidopsis thaliana] 6765 Seq. No. 7355 2.R1040 Contig ID $ncj7\overline{0}0980967.h1$ 5'-most EST 6766 Seq. No. 7356_1.R1040 Contig ID gsv701044894.h1 5'-most EST Method BLASTX NCBI GI g3695388 BLAST score 185 2.0e-13 E value 42 Match length 69 % identity (AF096371) No definition line found [Arabidopsis thaliana] NCBI Description 6767 Seq. No. 7356 2.R1040 Contig ID jex700907890.h15'-most EST 6768 Seq. No. Contig ID 7358 1.R1040 5'-most EST pxt700944366.hl Seq. No. 6769 Contig ID 7361 1.R1040 5'-most EST $leu7\overline{0}1148420.h1$ 6770 Seq. No. 7362 1.R1040 Contig ID 5'-most EST jex700909523.hl Method BLASTN g758249 NCBI GI BLAST score 427 E value 0.0e + 00Match length 543 95 % identity NCBI Description P.vulgaris mRNA for plasma membrane H+ ATPase Seq. No. 6771 7364 1.R1040 Contig ID 5'-most EST LIB3138-128-Q1-N1-A2

6772 Seq. No.

Contig ID 7366 1.R1040

5'-most EST LIB3030-002-Q1-B1-B9

Method BLASTX NCBI GI g2444178 BLAST score 509 E value 1.0e-51 Match length 160



```
% identity
                  (U94784) unconventional myosin [Helianthus annuus]
NCBI Description
                  6773
Seq. No.
Contig ID
                  7369 1.R1040
                  LIB3106-054-Q1-K1-E10
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2961346
BLAST score
                  387
E value
                  2.0e-37
Match length
                  96
                  70
% identity
                  (AL022140) pectinesterase like protein [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  6774
Contig ID
                  7370 1.R1040
                  trc700566871.h1
5'-most EST
Method
                  BLASTX
                  q173419
NCBI GI
BLAST score
                  253
E value
                  2.0e-36
Match length
                  118
                   67
% identity
                   (L11574) p68 RNA helicase [Schizosaccharomyces pombe]
NCBI Description
                  >qi 3810840 emb CAA21801 (AL032684) p68-like protein.
                   [Schizosaccharomyces pombe]
Seq. No.
                   6775
Contig ID
                  7376 1.R1040
                  LIB3030-001-Q1-B1-H4
5'-most EST
Method
                  BLASTX
                  q2459438
NCBI GI
                   248
BLAST score
                  5.0e-21
E value
Match length
                  133
% identity
                   51
                  (AC002332) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6776
                   7377 1.R1040
Contig ID
                  LIB3030-001-Q1-B1-H5
5'-most EST
                   6777
Seq. No.
                   7378 1.R1040
Contig ID
5'-most EST
                  LIB3092-044-Q1-K1-E3
                   BLASTX
Method
                   g2257756
NCBI GI
BLAST score
                   450
E value
                   3.0e-44
Match length
                   302
```

% identity NCBI Description

(U82815) nucleolar histone deacetylase HD2-p39 [Zea mays]

>gi 3650466 (AF026917) histone deacetylase HD2-p39 [Zea

mays]

39

Seq. No.



7378 2.R1040 Contig ID LIB3052-001-Q1-B1-H2 5'-most EST Method BLASTX q2257756 NCBI GI BLAST score 238 E value 1.0e-19 101 Match length % identity (U82815) nucleolar histone deacetylase HD2-p39 [Zea mays] NCBI Description >qi 3650466 (AF026917) histone deacetylase HD2-p39 [Zea mays] 6779 Seq. No.

7378 3.R1040 Contig ID uC-gmropic115a01b1 5'-most EST Method BLASTX NCBI GI q2257756 BLAST score 295

3.0e-26 E value 242 Match length 35 % identity

(U82815) nucleolar histone deacetylase HD2-p39 [Zea mays] NCBI Description

>gi 3650466 (AF026917) histone deacetylase HD2-p39 [Zea

mays]

6780 Seq. No.

7379 1.R1040 Contig ID

5'-most EST LIB3030-001-Q1-B1-H7

6781 Seq. No.

7380 1.R1040 Contig ID

jC-gmle01810084e03a1 5'-most EST

6782 Seq. No.

7380 2.R1040 Contig ID zhf700963876.h1 5'-most EST

6783 Seq. No.

7380 3.R1040 Contig ID 5'-most EST epx701105982.h1

Seq. No. 6784

7384 1.R1040 Contig ID 5'-most EST sat701004583.h1

Method BLASTX NCBI GI q1170508 BLAST score 2019 0.0e+00E value Match length 413 % identity 93 NCBI Description

EUKARYOTIC INITIATION FACTOR 4A-8 (EIF-4A-8)

>gi_2119931_pir__S60244 translation initiation factor

eIF-4A.8, anther-specific - common tobacco

 $>gi_475219_{emb}_CAA55639_(X79004)$ translation initiation

factor (eIF-4A) [Nicotiana tabacum]

>gi_475221_emb_CAA55640_ (X79005) translation initiation

factor (eIF-4A) [Nicotiana tabacum]



Seq. No. 6785

Contig ID 7384 3.R1040

5'-most EST LIB3\overline{1}06-106-Q1-K1-G7

Method BLASTX
NCBI GI g1170503
BLAST score 608
E value 4.0e-63
Match length 122
% identity 97

NCBI Description EUKARYOTIC INITIATION FACTOR 4A-1 (EIF-4A-1)

>gi_322503_pir__JC1452 translation initiation factor
eIF-4A1 - Arabidopsis thaliana >gi_16554_emb_CAA46188_
(X65052) eukaryotic translation initiation factor 4A-1

[Arabidopsis thaliana]

Seq. No. 6786

Contig ID 7384_4.R1040 5'-most EST trc700561210.h1

Seq. No. 6787

Contig ID 7384_5.R1040

5'-most EST uC-gmflminsoy043d12b1

Method BLASTX
NCBI GI g2500521
BLAST score 584
E value 2.0e-60
Match length 120
% identity 93

NCBI Description EUKARYOTIC INITIATION FACTOR 4A-15 (EIF-4A-15)

>qi 485945 emb CAA55739 (X79138) NeIF-4A15 [Nicotiana

tabacum]

Seq. No. 6788

Contig ID 7384 6.R1040

5'-most EST LIB3039-022-Q1-E1-H6

Method BLASTN
NCBI GI 9485986
BLAST score 108
E value 1.0e-53
Match length 224
% identity 87

NCBI Description N.tabacum NeIF-4A11 mRNA

Seq. No. 6789

Contig ID 7384_8.R1040 5'-most EST fde700876750.h1

Method BLASTN
NCBI GI g485944
BLAST score 69
E value 2.0e-30
Match length 157
% identity 86

NCBI Description N.tabacum NeIF-4A15 mRNA

Seq. No. 6790

Contig ID 7385_1.R1040



LIB3030-002-Q1-B1-A2 5'-most EST BLASTX Method g1935021 NCBI GI BLAST score 535 E value 1.0e-54 Match length 145 % identity 64 (Z93775) monosaccharid transport protein [Vicia faba] NCBI Description Seq. No. 7390 1.R1040 Contig ID yza700764039.hl 5'-most EST BLASTX Method q2245131 NCBI GI BLAST score 499 E value 6.0e-50 Match length 255 % identity 45 (Z97344) hypothetical protein [Arabidopsis thaliana] NCBI Description 6792 Seq. No. Contig ID 7391 1.R1040 LIB3087-002-Q1-K1-C6 5'-most EST BLASTX Method g3402701 NCBI GI 182 BLAST score E value 4.0e-13 75 Match length % identity (AC004261) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 6793 7393 1.R1040 Contig ID uC-gmflminsoy063d12b1 5'-most EST BLASTX Method g3047083 NCBI GI BLAST score 795 7.0e-85 E value 199 Match length % identity 84 (AF058914) similar to FLAP endonuclease-1 (SW:P39748) NCBI Description [Arabidopsis thaliana] 6794 Seq. No. 7393 2.R1040 Contig ID 5'-most EST $vwf7\overline{0}0674123.h1$ BLASTX Method g4454052 NCBI GI BLAST score 163 E value 8.0e-11 Match length 200 % identity 19 (AL035394) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 6795

Contig ID 7396 1.R1040 5'-most EST kl1701213446.h1



Seq. No. 6796

Contig ID 7397_1.R1040

5'-most EST LIB3030-001-Q1-B1-G12 Method BLASTX

NCBI GI g2129740
BLAST score 300
E value 8.0e-27
Match length 105
% identity 59

NCBI Description small nuclear ribonucleoprotein - Arabidopsis thaliana >gi 2129756 pir S71411 U1 snRNP 70K protein - Arabidopsis

thaliana >gi_1255711 (M93439) small nuclear

ribonucleoprotein [Arabidopsis thaliana] >gi_1354469 (U52909) U1 snRNP 70K protein [Arabidopsis thaliana]

Seq. No. 6797

Contig ID 7397_2.R1040

5'-most EST jC-gmro02910048h03a1

Method BLASTX
NCBI GI g2129740
BLAST score 388
E value 6.0e-38
Match length 88
% identity 81

NCBI Description small nuclear ribonucleoprotein - Arabidopsis thaliana

>gi 2129756_pir__S71411 U1 snRNP 70K protein - Arabidopsis

thaliana >qi 1255711 (M93439) small nuclear

ribonucleoprotein [Arabidopsis thaliana] >gi_1354469 (U52909) U1 snRNP 70K protein [Arabidopsis thaliana]

Seq. No. 6798

Contig ID 7398 1.R1040

5'-most EST fC-gmro7000748475f1

Method BLASTX
NCBI GI g3820531
BLAST score 592
E value 3.0e-61
Match length 169
% identity 63

NCBI Description (AF072736) beta-glucosidase [Pinus contorta]

Seq. No. 6799

Contig ID 7401_1.R1040 5'-most EST epx701106660.h1

Method BLASTX
NCBI GI g1946359
BLAST score 361
E value 2.0e-34
Match length 89
% identity 73

NCBI Description (U93215) unknown protein [Arabidopsis thaliana]

Seq. No. 6800

Contig ID 7403_1.R1040 5'-most EST hrw701060631.h1

Method BLASTN



```
NCBI GI g2274858
BLAST score 113
E value 2.0e-56
Match length 233
% identity 87
```

NCBI Description Arabidopsis thaliana mRNA for Cks1/Suc1 protein homologue

Seq. No. 6801

Contig ID 7404_1.R1040 5'-most EST ncj700987926.h1

Seq. No. 6802

Contig ID 7406_1.R1040 5'-most EST trc700565644.h1

Method BLASTX
NCBI GI g3004547
BLAST score 795
E value 1.0e-84
Match length 210
% identity 85

NCBI Description (AC003673) unknown protein [Arabidopsis thaliana] >gi 4185150 (AC005724) unknown protein [Arabidopsis

thaliana]

Seq. No. 6803

Contig ID 7408_1.R1040

5'-most EST uC-gmrominsoy071d11b1

Method BLASTX
NCBI GI g1894771
BLAST score 434
E value 3.0e-42
Match length 326
% identity 32

NCBI Description (Z92954) product highly similar to metabolite transport

proteins [Bacillus subtilis] >gi_2636109_emb_CAB15600.1_ (Z99122) similar to metabolite transport protein [Bacillus

subtilis]

Seq. No. 6804

Contig ID 7408_2.R1040

5'-most EST LIB3109-043-Q1-K1-F4

Seq. No. 6805

Contig ID 7409 1.R1040

5'-most EST LIB3\overline{1}39-072-P1-N1-F12

Method BLASTX
NCBI GI 9729668
BLAST score 397
E value 3.0e-38
Match length 140
% identity 59

NCBI Description HISTONE H1 >gi_2147479_pir__S65059 histone H1,

drought-inducible - Lycopersicon pennellii >gi_436823 (U01890) Solanum pennellii histone H1 [Solanum pennellii]

Seq. No. 6806

Contig ID 7409_2.R1040



```
LIB3170-067-Q1-K1-E8
5'-most EST
                  BLASTX
Method
                  g729668
NCBI GI
BLAST score
                  348
E value
                  2.0e-32
Match length
                  136
% identity
NCBI Description
```

HISTONE H1 >qi 2147479 pir S65059 histone H1, drought-inducible - Lycopersicon pennellii >gi 436823

(U01890) Solanum pennellii histone H1 [Solanum pennellii]

6807 Seq. No.

7410 1.R1040 Contig ID

LIB3106-062-Q1-K1-G9 5'-most EST

BLASTX Method NCBI GI q2642158 BLAST score 324 E value 1.0e-52 136 Match length % identity

NCBI Description (AC003000) hypothetical protein [Arabidopsis thaliana]

6808 Seq. No.

7410 2.R1040 Contig ID uaw700663338.h1 5'-most EST

Seq. No. 6809

7411 1.R1040 Contig ID

LIB3030-001-Q1-B1-E12 5'-most EST

Method BLASTX q2924509 NCBI GI BLAST score 666 E value 9.0e-70 Match length 229 % identity

(AL022023) subtilisin proteinase - like [Arabidopsis NCBI Description

thaliana]

6810 Seq. No.

7412 1.R1040 Contig ID vwf700676026.h1 5'-most EST

Seq. No. 6811

7413 1.R1040 Contig ID

LIB3030-001-Q1-B1-E3 5'-most EST

BLASTX Method g3738339 NCBI GI BLAST score 314 2.0e-28 E value 123 Match length % identity

(AC005170) putative kinase [Arabidopsis thaliana] NCBI Description

6812 Seq. No.

7413 2.R1040 Contig ID

LIB3049-006-Q1-E1-C6 5'-most EST



```
Seq. No.
                  7414 1.R1040
Contig ID
5'-most EST
                  dpv701097016.hl
                  BLASTX
Method
NCBI GI
                  g1076344
BLAST score
                  156
                   3.0e-10
E value
                  111
Match length
                   42
% identity
                  kinase-associated protein phosphatase precursor -
NCBI Description
                  Arabidopsis thaliana
                   6814
Seq. No.
                   7415 1.R1040
Contig ID
5'-most EST
                   jC-gmro02910064f09a1
Seq. No.
                   6815
Contig ID
                   7415 2.R1040
                   uC-qmflminsoy097f06b1
5'-most EST
                   6816
Seq. No.
Contig ID
                   7416 1.R1040
                   LIB3093-058-Q1-K1-D10
5'-most EST
                   BLASTX
Method
                   q4138583
NCBI GI
BLAST score
                   2250
E value
                   0.0e + 00
                   613
Match length
% identity
                  (Y10821) plastidic ATP/ADP-transporter [Solanum tuberosum]
NCBI Description
Seq. No.
                   6817
                   7416 2.R1040
Contig ID
                   LIB3092-002-Q1-K1-H1
5'-most EST
                   BLASTX
Method
                   g4138583
NCBI GI
BLAST score
                   346
                   2.0e-32
E value
                   103
Match length
% identity
                   71
                   (Y10821) plastidic ATP/ADP-transporter [Solanum tuberosum]
NCBI Description
                   6818
Seq. No.
                   7416 3.R1040
Contig ID
                   LIB3051-001-Q1-E1-A12
5'-most EST
                   BLASTX
Method
                   g4138583
NCBI GI
```

534 BLAST score 1.0e-54 E value Match length 117 % identity 91

NCBI Description

(Y10821) plastidic ATP/ADP-transporter [Solanum tuberosum]

Seq. No.

6819

Contig ID 5'-most EST 7418 1.R1040 sat701009519.h1



```
Seq. No.
                  7422 1.R1040
Contig ID
                  LIB3051-017-Q1-E1-H2
5'-most EST
                  BLASTX
Method
                  g3075391
NCBI GI
BLAST score
                  1716
                  0.0e + 00
E value
                  596
Match length
                   59
% identity
                  (AC004484) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7422 2.R1040
Contig ID
5'-most EST
                   jC-gmro02910010g05a1
                   6822
Seq. No.
                   7422 3.R1040
Contig ID
                   jC-gmle01810073c05a1
5'-most EST
                   BLASTX
Method
                   g3075391
NCBI GI
BLAST score
                   187
                   1.0e-13
E value
Match length
                   88
                   49
% identity
NCBI Description (AC004484) unknown protein [Arabidopsis thaliana]
Seq. No.
                   6823
                   7422 4.R1040
Contig ID
                   jC-qmf102220052b07a1
5'-most EST
                   BLASTX
Method
                   q3075391
NCBI GI
                   552
BLAST score
E value
                   9.0e-57
                   147
Match length
                   75
% identity
                   (AC004484) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   6824
                   7422 7.R1040
Contig ID
                   uC-gmropic088h02b1
5'-most EST
Seq. No.
                   6825
Contig ID
                   7428 1.R1040
                   LIB3109-005-Q1-K1-C10
5'-most EST
                   BLASTX
Method
                   g2342735
NCBI GI
                   463
BLAST score
                   4.0e-46
E value
                   130
Match length
                   75
% identity
                   (AC002341) unknown protein [Arabidopsis thaliana]
```

NCBI Description

6826 Seq. No.

7428 2.R1040 Contig ID kmv700738733.h1 5'-most EST

BLASTX Method NCBI GI g2342735



```
227
BLAST score
                  9.0e-19
E value
                  57
Match length
                  81
% identity
NCBI Description (AC002341) unknown protein [Arabidopsis thaliana]
                  6827
Seq. No.
                  7430 1.R1040
Contig ID
                  jex700904532.hl
5'-most EST
                   6828
Seq. No.
                  7433 1.R1040
Contig ID
                  g4290321
5'-most EST
                   6829
Seq. No.
                   7442 1.R1040
Contig ID
                   LIB3040-040-Q1-E1-B3
5'-most EST
                   {\tt BLASTX}
Method
                   g2135841
NCBI GI
BLAST score
                   184
                   7.0e-13
E value
                   78
Match length
                   46
% identity
                   nucleic acid-binding protein E5.1 - human >gi 598231
NCBI Description
                   (L37368) RNA-binding protein [Homo sapiens]
                   6830
Seq. No.
                   7448 1.R1040
Contig ID
                   LIB3030-001-Q1-B1-B5
5'-most EST
                   BLASTX
Method
                   q3915022
NCBI GI
                   768
BLAST score
                   5.0e-82
E value
Match length
                   178
% identity
                   83
                   SUCROSE-PHOSPHATE SYNTHASE 2
NCBI Description
                   (UDP-GLUCOSE-FRUCTOSE-PHOSPHATE GLUCOSYLTRANSFERASE 2)
                   >gi_2190350_emb_CAA72491 (Y11795) sucrose-phosphate
                   synthase [Craterostigma plantagineum]
                   6831
Seq. No.
                   7449 1.R1040
Contig ID
5'-most EST
                   g5606523
                   6832
Seq. No.
                   7450 1.R1040
Contig ID
                   LIB3106-101-Q1-K1-F9
5'-most EST
Method
                   BLASTX
                   g3915665
NCBI GI
```

NCBI GI g391566 BLAST score 236 E value 3.0e-27 Match length 143 % identity 44

NCBI Description ELONGATION FACTOR G (EF-G) >gi_2688449 (AE001155) translation elongation factor G (fus-1) [Borrelia

burgdorferi]



```
Seq. No.
                     7456 1.R1040
 Contig ID
                     uxk700667188.hl
 5'-most EST
                     BLASTX
 Method
                     q4490733
 NCBI GI
 BLAST score
                     574
                     5.0e-59
 E value
                     249
 Match length
 % identity
                     (AL035709) putative protein [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
                     7458 1.R1040
 Contig ID
                     zsg701119169.h1
 5'-most EST
 Seq. No.
                      7459 1.R1040
 Contig ID
                      jC-qmf102220079b07a1
 5'-most EST
                      6836
 Seq. No.
                      7465 1.R1040
 Contig ID
 5'-most EST
                      fde700875880.hl
                      BLASTX
 Method
 NCBI GI
                      q1169586
                      582
 BLAST score
                      4.0e-60
 E value
 Match length
                      127
 % identity
                      FRUCTOSE-1, 6-BISPHOSPHATASE, CYTOSOLIC
 NCBI Description
                      (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)
                      (CY-F1) >gi_542079_pir_ S41287 fructose-bisphosphatase (EC 3.1.3.11) - potato >gi_440591_emb_CAA54265_ (X76946)
                      fructose-1,6-bisphosphatase [Solanum tuberosum]
                      6837
  Seq. No.
                      7465 2.R1040
  Contig ID
                      LIB3106-095-Q1-K1-H11
  5'-most EST
                      BLASTX
Method
                      g1169586
  NCBI GI
                      273
  BLAST score
                      4.0e-24
  E value
  Match length
                      57
                      89
  % identity
                      FRUCTOSE-1, 6-BISPHOSPHATASE, CYTOSOLIC
  NCBI Description
                      (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)
                      (CY-F1) >gi_542079_pir__S41287 fructose-bisphosphatase (EC 3.1.3.11) - potato >gi_440591_emb_CAA54265_ (X76946)
                      fructose-1,6-bisphosphatase [Solanum tuberosum]
                      6838
  Seq. No.
                      7466 1.R1040
  Contig ID
                      smc7\overline{0}0747854.h1
  5'-most EST
```

BLASTX Method q3107931 NCBI GI 1876 BLAST score 0.0e+00E value Match length 461



% identity 83
NCBI Description (AB013353) UDP-glucose pyrophosphorylase [Pyrus pyrifolia]

Seq. No. 6839

Contig ID 7466_2.R1040 5'-most EST pmv700893276.h1

Method BLASTN
NCBI GI g3107930
BLAST score 99
E value 4.0e-48
Match length 316
% identity 83

NCBI Description Pyrus pyrifolia mRNA for UDP-glucose pyrophosphorylase,

complete cds

Seq. No. 6840

Contig ID 7466_3.R1040

5'-most EST jC-gmst02400062h08d1

Method BLASTX
NCBI GI g136739
BLAST score 232
E value 3.0e-19
Match length 59
% identity 75

NCBI Description UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE (UDP-GLUCOSE

PYROPHOSPHORYLASE) (UDPGP) >gi_67061_pir__XNPOU

UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) -

potato >gi_218001_dbj_BAA00570_ (D00667) UDP-glucose

pyrophosphorylase precursor [Solanum tuberosum]

Seq. No. 6841

Contig ID 7466_5.R1040 5'-most EST pxt700946288.h1

Method BLASTX
NCBI GI g3107931
BLAST score 175
E value 1.0e-12
Match length 43
% identity 72

NCBI Description (AB013353) UDP-glucose pyrophosphorylase [Pyrus pyrifolia]

Seq. No. 6842

Contig ID 7466_9.R1040 5'-most EST ncj700981983.h1

Method BLASTX
NCBI GI g136739
BLAST score 123
E value 3.0e-14
Match length 55
% identity 80

NCBI Description UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE (UDP-GLUCOSE

PYROPHOSPHORYLASE) (UDPGP) >gi_67061_pir__XNPOU

UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) -

potato >gi 218001 dbj_BAA00570 (D00667) UDP-glucose

pyrophosphorylase precursor [Solanum tuberosum]

Seq. No. 6843



Contig ID 7467_1.R1040 5'-most EST leu701153852.h1 Method BLASTX

NCBI GI g461498
BLAST score 1244
E value 1.0e-137
Match length 293
% identity 80

NCBI Description ALANINE AMINOTRANSFERASE 2 (GPT) (GLUTAMIC--PYRUVIC TRANSAMINASE 2) (GLUTAMIC--ALANINE TRANSAMINASE 2)

(ALAAT-2) >gi_320619_pir__S28429 alanine transaminase (EC 2.6.1.2) - proso millet >gi_296204_emb_CAA49199_ (X69421)

alanine aminotransferase [Panicum miliaceum]

Seq. No. 6844

Contig ID 7467_4.R1040 5'-most EST kmv700740452.h1

Method BLASTX
NCBI GI 9461498
BLAST score 184
E value 8.0e-14
Match length 47
% identity 72

NCBI Description ALANINE AMINOTRANSFERASE 2 (GPT) (GLUTAMIC--PYRUVIC TRANSAMINASE 2) (GLUTAMIC--ALANINE TRANSAMINASE 2)

(ALAAT-2) >gi_320619_pir__S28429 alanine transaminase (EC 2.6.1.2) - proso millet >gi_296204_emb_CAA49199_ (X69421)

alanine aminotransferase [Panicum miliaceum]

Seq. No. 6845

Contig ID 7470_1.R1040

5'-most EST LIB3030-005-Q1-B1-E5

Method BLASTX
NCBI GI g3914940
BLAST score 1579
E value 1.0e-176
Match length 378
% identity 80

NCBI Description SEDOHEPTULOSE-1,7-BISPHOSPHATASE CHLOROPLAST PRECURSOR

(SEDOHEPTULOSE-BISPHOSPHATASE) (SBPASE) (SED(1,7)P2ASE) >qi 2529376 (L76556) sedoheptulose-1,7-bisphosphatase

[Spinacia oleracea]

Seq. No. 6846

Contig ID 7470_2.R1040

5'-most EST LIB3092-013-Q1-K1-D11

Seq. No. 6847

Contig ID 7474_1.R1040 5'-most EST fua701041768.h1

Method BLASTX
NCBI GI g2492952
BLAST score 1304
E value 1.0e-144
Match length 312
% identity 78

NCBI Description CHORISMATE SYNTHASE 1 PRECURSOR



(5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE PHOSPHOLYASE 1)
>gi_542026_pir__S40410 chorismate synthase (EC 4.6.1.4) 1
precursor - tomato >gi_410482_emb_CAA79859_ (Z21796)
chorismate synthase 1 [Lycopersicon esculentum]

Seq. No. 6848

Contig ID 7477_1.R1040

5'-most EST jC-gmf102220131g06a1

Method BLASTX
NCBI GI g1871192
BLAST score 1127
E value 1.0e-147
Match length 485
% identity 61

NCBI Description (U90439) Cys3His zinc finger protein isolog [Arabidopsis

thaliana]

Seq. No. 6849

Contig ID 7478_1.R1040

5'-most EST LIB3139-045-P1-N1-B6

Method BLASTX
NCBI GI g1763063
BLAST score 328
E value 4.0e-30
Match length 178
% identity 43

NCBI Description (U68763) SCOF-1 [Glycine max]

Seq. No. 6850

Contig ID 7478_2.R1040

5'-most EST LIB3092-049-Q1-K1-G5

Method BLASTX
NCBI GI g1763063
BLAST score 272
E value 8.0e-24
Match length 112
% identity 58

NCBI Description (U68763) SCOF-1 [Glycine max]

Seq. No. 6851

Contig ID 7478_3.R1040

5'-most EST LIB3107-036-Q1-K1-B11

Method BLASTX
NCBI GI g439489
BLAST score 219
E value 9.0e-18
Match length 113
% identity 35

NCBI Description (D26084) zinc-finger DNA binding protein [Petunia x

hybrida]

Seq. No. 6852

Contig ID 7478_4.R1040

5'-most EST LIB3050-022-Q1-K1-F11

Seq. No. 6853

Contig ID 7479_1.R1040

Match length

% identity

675

90



uC-gmrominsoy040h09b1 5'-most EST BLASTX Method g100226 NCBI GI BLAST score 332 E value 6.0e-31 Match length 143 % identity 51 hypothetical protein - tomato >gi_19275_emb_CAA78112_ NCBI Description (Z12127) protein of unknown function [Lycopersicon esculentum] >gi 445619 prf 1909366A Leu zipper protein [Lycopersicon esculentum] 6854 Seq. No. 7481 1.R1040 Contig ID 5'-most EST zsg701127164.hl Method BLASTX g4433048 NCBI GI BLAST score 635 E value 9.0e-66 Match length 146 % identity 84 (D26578) DNA-binding protein [Daucus carota] NCBI Description 6855 Seq. No. 7481 2.R1040 Contig ID 5'-most EST LIB3065-014-Q1-N1-F3 Method BLASTX g1168547 NCBI GI BLAST score 176 3.0e-12 E value Match length 188 33 % identity HOMEOBOX-LEUCINE ZIPPER PROTEIN ATHB-6 (HD-ZIP PROTEIN NCBI Description ATHB-6) >gi 629505 pir S47136 homeotic protein Athb-6 -Arabidopsis thaliana >gi_499162_emb_CAA47427_ (X67034) Athb-6 [Arabidopsis thaliana] Seq. No. 6856 7481 3.R1040 Contig ID 5'-most EST LIB3050-023-Q1-K1-B2 Method BLASTX NCBI GI q4433048 274 BLAST score 6.0e-24 E value Match length 68 % identity 79 NCBI Description (D26578) DNA-binding protein [Daucus carota] Seq. No. 6857 7483 1.R1040 Contig ID 5'-most EST sat701010763.hl Method BLASTN NCBI GI q2318116 BLAST score 392 E value 0.0e+00

Seq. No.

Contig ID

6864

7501 1.R1040



```
NCBI Description Pisum sativum Mg-chelatase subunit D (ChlD) mRNA, complete
                  6858
Seq. No.
Contig ID
                  7491 1.R1040
                  LIB3030-001-Q1-B1-H10
5'-most EST
                  BLASTX
Method
                  q4468990
NCBI GI
BLAST score
                  496
                  7.0e-50
E value
                  245
Match length
                  45
% identity
NCBI Description (AL035605) putative protein [Arabidopsis thaliana]
Seq. No.
                  6859
                  7492 1.R1040
Contig ID
5'-most EST
                  LIB3030-002-Q1-B1-B3
Method
                  BLASTX
                  g2688830
NCBI GI
BLAST score
                  180
E value
                  5.0e-13
Match length
                  52
% identity
                  65
                  (AF000952) putative sugar transporter [Prunus armeniaca]
NCBI Description
                  6860
Seq. No.
                  7497 1.R1040
Contig ID
                  LIB3030-003-Q1-B1-G3
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1213629
BLAST score
                  367
                  5.0e-35
E value
                  113
Match length
% identity
                  60
NCBI Description (X95991) pectinesterase [Prunus persica]
                  6861
Seq. No.
                  7497 2.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy041b10b1
Method
                  BLASTX
                  a1213629
NCBI GI
BLAST score
                  164
                  2.0e-11
E value
                  82
Match length
% identity
                  39
NCBI Description
                  (X95991) pectinesterase [Prunus persica]
Seq. No.
                  6862
                  7499 1.R1040
Contig ID
5'-most EST
                  bth700846254.h1
Seq. No.
                  6863
                  7500 1.R1040
Contig ID
5'-most EST
                  jC-gmf102220073a09d1
```



```
5'-most EST
                   g5753413
                   BLASTX
Method
                  · q2129564
NCBI GI
                   2107
BLAST score
                   0.0e + 00
E value
Match length
                   585
% identity
```

NCBI Description cryptochrome 2 apoprotein - Arabidopsis thaliana

6865 Seq. No.

7501 2.R1040 Contig ID

LIB3094-102-Q1-K1-H6 5'-most EST

Method BLASTX q730324 NCBI GI 276 BLAST score E value 2.0e-24 Match length 55 % identity

DEOXYRIBODIPYRIMIDINE PHOTOLYASE (DNA PHOTOLYASE) NCBI Description

(PHOTOREACTIVATING ENZYME) >gi 1084356 pir S48120 deoxyribodipyrimidine photo-lyase (EC $\overline{4.1.99.3}$) - white mustard >gi_414196_emb_CAA50898_ (X72019) photolyase

[Sinapis alba]

6866 Seq. No.

7501 3.R1040 Contig ID

5'-most EST jC-qmle01810030c10a2

BLASTX Method NCBI GI q730324 BLAST score 161 1.0e-10 E value Match length 37 70 % identity

NCBI Description

DEOXYRIBODIPYRIMIDINE PHOTOLYASE (DNA PHOTOLYASE) (PHOTOREACTIVATING ENZYME) >gi_1084356_pir__S48120 deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - white

mustard >gi 414196 emb CAA50898 (X72019) photolyase

[Sinapis alba]

6867 Seq. No.

7501 4.R1040 Contig ID

LIB3030-004-Q1-B1-B2 5'-most EST

BLASTX Method g2129564 NCBI GI BLAST score 219 2.0e-17 E value 170 Match length 39 % identity

NCBI Description cryptochrome 2 apoprotein - Arabidopsis thaliana

Seq. No. 6868

7501 5.R1040 Contig ID kl1701203603.h2 5'-most EST

6869 Seq. No.

Contig ID 7502 1.R1040

5'-most EST jC-qmro02800040g05d1



```
6870
Seq. No.
                  7504 1.R1040
Contig ID
5'-most EST
                  LIB3030-004-Q1-B1-G3
Method
                  BLASTX
NCBI GI
                  q2769566
BLAST score
                   494
                   1.0e-49
E value
                   156
Match length
% identity
                   (Y10477) chloroplast thylakoidal processing peptidase
NCBI Description
                   [Arabidopsis thaliana]
                   6871
Seq. No.
Contig ID
                   7506 1.R1040
5'-most EST
                   uC-gmronoir038a04b1
Method
                   BLASTX
NCBI GI
                   q3819710
BLAST score
                   265
                   6.0e-23
E value
                   60
Match length
% identity
                   80
                   (AJ224161) delta-8 sphingolipid desaturase [Arabidopsis
NCBI Description
                   thaliana]
                   6872
Seq. No.
                   7506 4.R1040
Contig ID
5'-most EST
                   zzp700834013.h1
Method
                   BLASTX
                   g3702328
NCBI GI
BLAST score
                   309
                   2.0e-28
E value
Match length
                   70
                   77
% identity
                   (ACO05397) putative cytochrome b5 [Arabidopsis thaliana]
NCBI Description
                   6873
Seq. No.
                   7508 1.R1040
Contig ID
                   LIB3138-042-Q1-N1-A5
5'-most EST
Seq. No.
                   6874
                   7508 2.R1040
Contig ID
                   zzp700833303.h1
5'-most EST
                   BLASTN
Method
                   g1750375
NCBI GI
BLAST score
                   65
                   8.0e-28
E value
                   161
Match length
                   85
% identity
                   Arabidopsis thaliana ubiquitin activating enzyme (UBA1)
NCBI Description
                   gene, complete cds
                   6875
```

Seq. No.

7511 1.R1040 Contig ID 5'-most EST kl1701212848.hl

6876 Seq. No.

#1 (PM)



7514 1.R1040 Contig ID jC-gmro02910061a04a1 5'-most EST BLASTX Method NCBI GI q4314365 191 BLAST score 3.0e-14 E value Match length 142 39 % identity (AC006340) putative copia-like retrotransposon Hopscotch NCBI Description [Arabidopsis thaliana] 6877 Seq. No. 7516 1.R1040 Contig ID LIB3030-007-Q1-B1-A9 5'-most EST Method BLASTX NCBI GI g4337175

BLAST score 501 2.0e-50 E value Match length 161 % identity 60

NCBI Description

(AC006416) ESTs gb_T20589, gb_T04648, gb_AA597906, gb_T04111, gb_R84180, gb_R65428, gb_T44439, gb_T76570, gb_R90004, gb_T45020, gb_T42457, gb_T20921, gb_AA042762 and gb AA720210 come from this gene. [Arabidopsis thaliana]

6878 Seq. No.

7518 1.R1040 Contig ID

5'-most EST LIB3051-026-Q1-K1-F9

Method BLASTX q2689720 NCBI GI 249 BLAST score 3.0e-21 E value 83 Match length 60 % identity

(AF037168) DnaJ homologue [Arabidopsis thaliana] NCBI Description

Seq. No.

6879

7519 1.R1040 Contig ID

5'-most EST LIB3030-007-Q1-B1-G4

Seq. No.

6880

7520 1.R1040 Contig ID

LIB3106-112-Q1-K1-C7 5'-most EST

Seq. No.

6881

Contig ID 7522 1.R1040

LIB3030-008-Q1-B1-A9 5'-most EST

Seq. No.

6882

7523 1.R1040 Contig ID

5'-most EST LIB3030-008-Q1-B1-B4

Seq. No.

6883

7525 1.R1040 Contig ID

5'-most EST jC-gmst02400055g12d1

Seq. No. 6884



```
7527 1.R1040
Contig ID
                  LIB3030-008-Q1-B1-G2
5'-most EST
                  BLASTX
Method
                  g3757523
NCBI GI
BLAST score
                  894
E value
                  1.0e-96
                  225
Match length
% identity
NCBI Description (AC005167) putative transportin [Arabidopsis thaliana]
                   6885
Seq. No.
                   7530 1.R1040
Contig ID
                   uC-gmropic096d09b1
5'-most EST
                   BLASTX
Method
                   q3033374
NCBI GI
BLAST score
                   192
                   1.0e-14
E value
Match length
                   137
                   39
% identity
NCBI Description (AC004238) unknown protein [Arabidopsis thaliana]
Seq. No.
                   7532 1.R1040
Contig ID
                   uC-gmronoir044e01b1
5'-most EST
Seq. No.
                   7532 2.R1040
Contig ID
                   uC-gmflminsoy032g03b1
5'-most EST
Seq. No.
                   6888
                   7535 1.R1040
Contig ID
                   fC-gmst700664572d3
5'-most EST
                   BLASTX
Method
                   g3925363
NCBI GI
                   1011
BLAST score
                   1.0e-110
E value
Match length
                   267
                   78
% identity
                   (AF067961) homeodomain protein [Malus domestica]
NCBI Description
                   6889
Seq. No.
                   7535 2.R1040
Contig ID
                   fC-gmst700664572d5
 5'-most EST
                   BLASTX
Method
                   g3925363
NCBI GI
                   670
BLAST score
                   3.0e-70
E value
                   161
Match length
                   86
 % identity
                   (AF067961) homeodomain protein [Malus domestica]
NCBI Description
 Seq. No.
                   6890
```

Contig ID 7536_1.R1040

5'-most EST LIB3030-009-Q1-B1-C10

Method BLASTX
NCBI GI g3047124
BLAST score 191



```
2.0e-14
E value
                   62
Match length
% identity
                   (AF058919) No definition line found [Arabidopsis thaliana]
NCBI Description
                   6891
Seq. No.
                   7538 1.R1040
Contig ID
                   fC-qmro700797913a1
5'-most EST
                   BLASTX
Method
                   q2627181
NCBI GI
                   682
BLAST score
                   6.0e-72
E value
                   153
Match length
                   80
% identity
                  (D89619) cycloartenol synthase [Pisum sativum]
NCBI Description
                   6892
Seq. No.
                   7540 1.R1040
Contig ID
                   epx701107112.h1
5'-most EST
                   BLASTX
Method
                   q4454471
NCBI GI
                   685
BLAST score
                   8.0e-72
E value
                   315
Match length
                   48
% identity
                   (AC006234) putative G protein coupled receptor [Arabidopsis
NCBI Description
                   thaliana]
                   6893
Seq. No.
                   7541 1.R1040
Contig ID
                   LIB3030-009-Q1-B1-C4
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3522950
BLAST score
                   228
                   1.0e-18
E value
                    99
Match length
                    49
 % identity
                   (AC004411) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                    6894
 Seq. No.
                    7545 1.R1040
 Contig ID
                    LIB3030-009-Q1-B1-E1
 5'-most EST
Seq. No.
                    6895
                    7551 1.R1040
 Contig ID
                    k117\overline{0}1207181.h1
 5'-most EST
                    BLASTN
 Method
                    g1044939
 NCBI GI
                    226
 BLAST score
                    1.0e-124
 E value
                    432
 Match length
                    91
 % identity
                   Z.mays mRNA for ubiquitin/ribosomal protein S27a fusion
 NCBI Description
                    protein
```

7551_3.R1040

6896

Seq. No.

Contig ID



```
taw700660412.hl
5'-most EST
                  BLASTN
Method
                  q967984
NCBI GI
                  142
BLAST score
                  7.0e-74
E value
Match length
                  285
% identity
                  Oryza sativa (clone rma630) ribosomal protein-linked
NCBI Description
                  ubiquitin mRNA, complete cds
                  6897
Seq. No.
                  7551 6.R1040
Contig ID
                  jC-gmst02400060a06d1
5'-most EST
                  BLASTN
Method
                  q967984
NCBI GI
BLAST score
                  126
E value
                  2.0e-64
Match length
                  221
% identity
                  Oryza sativa (clone rma630) ribosomal protein-linked
NCBI Description
                  ubiquitin mRNA, complete cds
                   6898
Seq. No.
Contig ID
                   7552 1.R1040
                  LIB3055-004-Q1-N1-G8
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2500079
                   166
BLAST score
                   2.0e-16
E value
Match length
                   296
                   13
% identity
                   X-LINKED RETINITIS PIGMENTOSA GTPASE REGULATOR >gi_1314871
NCBI Description
                   (U57629) retinitis pigmentosa GTPase regulator [Homo
                   sapiens] >gi_2204218_emb_CAA66258_ (X97668) XLRP3 [Homo
                   sapiens] >gi_4506581_ref_NP_000319.1_pRP3_ retinitis
                   pigmentosa 3 (X-linked recessive)
                   6899
Seq. No.
                   7552 2.R1040
Contig ID
                   fde700876589.hl
5'-most EST
                   6900
Seq. No.
                   7552 3.R1040
Contig ID
                   LIB3138-090-P1-N1-B6
5'-most EST
                   BLASTX
Method
                   g1477565
NCBI GI
                   162
BLAST score
                   6.0e-11
E value
                   112
Match length
 % identity
NCBI Description
                   (U50078) p532 [Homo sapiens]
                   >gi 4557026 ref NP 003913.1 pHERC1 hect (homologous to the
                   E6-AP (UBE3A) carboxyl terminus) domain and RCC1
```

1282

(CHC1)-like domain (RLD)

6901

7556_1.R1040

Seq. No.

Contig ID



```
LIB3030-009-Q1-B1-G9
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2244856
BLAST score
                  292
E value
                  4.0e-26
Match length
                  128
% identity
NCBI Description (Z97337) unnamed protein product [Arabidopsis thaliana]
Seq. No.
                  7556 2.R1040
Contig ID
5'-most EST
                  ssr700556456.hl
                  6903
Seq. No.
                  7557 1.R1040
Contig ID
5'-most EST
                  LIB3107-006-Q1-K1-B11
Method
                  BLASTX
NCBI GI
                  q3860333
BLAST score
                  449
                  3.0e-44
E value
                  124
Match length
% identity
                  69
NCBI Description (AJ012693) basic blue copper protein [Cicer arietinum]
                   6904
Seq. No.
Contig ID
                  7557 2.R1040
                  leu701152331.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3860333
BLAST score
                   432
                   2.0e-42
E value
Match length
                  127
                   65
% identity
                  (AJ012693) basic blue copper protein [Cicer arietinum]
NCBI Description
                   6905
Seq. No.
                   7558 1.R1040
Contig ID
5'-most EST
                   uC-gmflminsoy028b03b1
                   6906
Seq. No.
Contig ID
                   7562 1.R1040
                   LIB3030-010-Q1-B1-A12
5'-most EST
                   6907
Seq. No.
                   7562 2.R1040
Contig ID
                   LIB3094-074-Q1-K1-H5
5'-most EST
                   6908
Seq. No.
                   7565 1.R1040
Contig ID
                   LIB3030-010-Q1-B1-C12
5'-most EST
Method
                   BLASTX
                   q2191175
NCBI GI
BLAST score
                   472
E value
                   7.0e-47
Match length
                   264
% identity
                   47
```

NCBI Description (AF007270) A IG002P16.24 gene product [Arabidopsis



thaliana]

Seq. No. 6909

Contig ID 7565 2.R1040

5'-most EST jC-gmf102220056e09a1

Seq. No. 6910

Contig ID 7566 1.R1040

5'-most EST uC-qmronoir047b10b1

Seq. No. 6911

Contig ID 7567_1.R1040

5'-most EST LIB3053-006-Q1-N1-G4

Method BLASTX
NCBI GI g3721926
BLAST score 2332
E value 0.0e+00
Match length 543
% identity 87

NCBI Description (AB017480) chloroplast FtsH protease [Nicotiana tabacum]

Seq. No. 6912

Contig ID 7567_2.R1040 5'-most EST rca700998469.h1

Method BLASTX
NCBI GI g2492515
BLAST score 540

BLAST score 540 E value 6.0e-55 Match length 191 % identity 61

NCBI Description CELL DIVISION PROTEIN FTSH HOMOLOG PRECURSOR

>gi_2129924_pir__S58298 ATPase - pepper (fragment)

>gi_929013_emb_CAA62084_ (X90472) ATPase [Capsicum annuum]

Seq. No. 6913

Contig ID 7570_1.R1040

5'-most EST LIB3030-010-Q1-B1-D9

Seq. No. 6914

Contig ID 7575 1.R1040 5'-most EST zzp700836238.h1

Seq. No. 6915

Contig ID 7576_1.R1040 5'-most EST zsg701130230.h1

Seq. No. 6916

Contig ID 7577_1.R1040

5'-most EST LIB3030-010-Q1-B1-F12

Seq. No. 6917

Contig ID 7581_1.R1040

5'-most EST LIB3030-010-Q1-B1-G11

Method BLASTX
NCBI GI g2245108
BLAST score 594
E value 3.0e-61



Match length 203 % identity 57

NCBI Description (Z97343) EREBP-4 homolog [Arabidopsis thaliana]

Seq. No. 6918

Contig ID 7583_1.R1040

5'-most EST LIB3040-014-Q1-E1-E10

Method BLASTX
NCBI GI g3122386
BLAST score 1798
E value 0.0e+00
Match length 356
% identity 91

NCBI Description WD-40 REPEAT PROTEIN MSI1 >gi_2394227 (AF016845) WD-40

repeat protein [Lycopersicon esculentum]

Seq. No. 6919

Contig ID 7583 2.R1040 5'-most EST epx701106768.h1

Seq. No. 6920

Contig ID 7583_3.R1040 5'-most EST hrw701063160.h1

Seq. No. 6921

Contig ID 7583_4.R1040

5'-most EST LIB3109-029-Q1-K1-G10

Seq. No. 6922

Contig ID 7586_1.R1040

5'-most EST LIB3030-012-Q1-B1-A2

Method BLASTX
NCBI GI 94218120
BLAST score 757
E value 1.0e-85
Match length 219
% identity 70

NCBI Description (AL035353) Proline-rich APG-like protein [Arabidopsis

thaliana]

Seq. No. 6923

Contig ID 7586_2.R1040 5'-most EST wrg700790986.h1

Method BLASTX
NCBI GI g4218120
BLAST score 321
E value 7.0e-30
Match length 84
% identity 70

NCBI Description (AL035353) Proline-rich APG-like protein [Arabidopsis

thaliana]

Seq. No. 6924

Contig ID 7591_1.R1040

5'-most EST LIB3030-012-Q1-B1-D1

Seq. No. 6925



```
7592 1.R1040
Contig ID
                   q560<del>6</del>539
5'-most EST
                   BLASTX
Method
                   g4510373
NCBI GI
                   623
BLAST score
                   2.0e-64
E value
Match length
                   171
                   64
% identity
                   (AC007017) putative harpin-induced protein [Arabidopsis
NCBI Description
                   thaliana]
                   6926
Seq. No.
                   7592 2.R1040
Contig ID
                   qsv701052671.h1
5'-most EST
                   BLASTX
Method
                   g4510373
NCBI GI
BLAST score
                   241
                   5.0e-20
E value
                   92
Match length
% identity
                   (AC007017) putative harpin-induced protein [Arabidopsis
NCBI Description
                   thaliana]
                   6927
Seq. No.
                   7595 1.R1040
Contig ID
                   jC-gmf102220062e05a1
5'-most EST
                   BLASTX
Method
                   q1255951
NCBI GI
                   1926
BLAST score
                   0.0e + 00
E value
                   515
Match length
% identity
                    69
                   (X96932) PS60 [Nicotiana tabacum]
NCBI Description
                    6928
Seq. No.
                    7598 1.R1040
Contig ID
                    fC-gmf1700904909f4
 5'-most EST
Method
                    BLASTX
                    q4098246
NCBI GI
                    1035
 BLAST score
 E value
                    1.0e-113
 Match length
                    282
 % identity
                    86
                    (U76410) homeobox 2 protein [Lycopersicon esculentum]
 NCBI Description
                    6929
 Seq. No.
                    7598 2.R1040
 Contig ID
                    jC-gmst02400067a07a1
 5'-most EST
                    BLASTX
 Method
                    g1170191
 NCBI GI
                    583
 BLAST score
                    3.0e-60
 E value
                    150
 Match length
                    77
 % identity
                    HOMEOBOX PROTEIN HD1 >gi_1076449_pir__S47535
 NCBI Description
```

1286

>gi_453949_emb_CAA82314_ (Z29073) homeodomain-containing

homeodomain-containing protein - rape



protein [Brassica napus] >gi_1090522_prf__2019252A homeobox protein [Brassica napus]

6930 Seq. No.

7598 3.R1040 Contig ID uC-gmropic045e12b1 5'-most EST

BLASTX Method NCBI GI g4098246 479 BLAST score 4.0e-48 E value 118 Match length 79 % identity

(U76410) homeobox 2 protein [Lycopersicon esculentum] NCBI Description

6931 Seq. No.

7598 4.R1040 Contig ID wrg700790517.h2 5'-most EST

BLASTX Method g4098246 NCBI GI 359 BLAST score 5.0e-34 E value 66 Match length 97 % identity

(U76410) homeobox 2 protein [Lycopersicon esculentum] NCBI Description

6932 Seq. No.

7598 5.R1040 Contig ID leu701145388.h1 5'-most EST

BLASTX Method g1170191 NCBI GI 245 BLAST score 3.0e-21 E value 78 Match length 67 % identity

HOMEOBOX PROTEIN HD1 >gi_1076449_pir__S47535 NCBI Description

homeodomain-containing protein - rape

>gi_453949_emb_CAA82314 (Z29073) homeodomain-containing
protein [Brassica napus] >gi_1090522_prf__2019252A homeobox

protein [Brassica napus]

6933 Seq. No.

7604 1.R1040 Contig ID leu701155424.h1 5'-most EST

6934 Seq. No.

7607 1.R1040 Contia ID $k117\overline{0}1202720.h1$ 5'-most EST

Method BLASTN g2760170 NCBI GI 34 BLAST score 2.0e-09 E value 97 Match length 89 % identity

Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description

MIO24, complete sequence [Arabidopsis thaliana]

6935 Seq. No.

% identity



```
7607 2.R1040
Contig ID
                   zhf700962831.h1
5'-most EST
                   6936
Seq. No.
                   7611 1.R1040
Contig ID
                  LIB3170-003-Q1-K1-F8
5'-most EST
                   BLASTN
Method
NCBI GI
                   g3869254
                   152
BLAST score
                   1.0e-79
E value
                   328
Match length
                   87
% identity
                   Pisum sativum defender against death homolog Peadad
NCBI Description
                   (PEADAD) mRNA, complete cds
                   6937
Seq. No.
                   7612 1.R1040
Contig ID
                   LIB3170-056-Q1-K1-C12
5'-most EST
                   6938
Seq. No.
                   7616 1.R1040
Contig ID
                   LIB3170-001-Q1-K1-E10
5'-most EST
                   BLASTN
Method
                   q498788
NCBI GI
                   195
BLAST score
                   1.0e-105
E value
                   499
Match length
                   85
% identity
                   S.tuberosum (Desiree) cycl II mRNA. for cytochrome c1
NCBI Description
                   6939
Seq. No.
                   7616 2.R1040
Contig ID
                   uC-qmflminsoy035g09b1
5'-most EST
                   BLASTX
Method
                   q440953
NCBI GI
                   418
BLAST score
                   5.0e-41
E value
Match length
                   93
% identity
                    (S66866) cytochrome c1 [Solanum tuberosum=potatoes,
NCBI Description
                   Desiree, monohaploid line AM 80/5793, Peptide
                   Mitochondrial, 320 aa] [Solanum tuberosum]
Seq. No.
                    6940
                    7616 3.R1040
Contig ID
                    zhf7\overline{0}0963832.h1
5'-most EST
Seq. No.
                    6941
                    7616 4.R1040
 Contig ID
                    jex700906970.h1
 5'-most EST
                    BLASTN
Method
                    g498788
NCBI GI
                    48
 BLAST score
                    7.0e-18
 E value
                    108
 Match length
                    86
```

NCBI Description S.tuberosum (Desiree) cycl II mRNA. for cytochrome c1



Seq. No. 6942

Contig ID 7621 1.R1040

5'-most EST LIB3\overline{107}-056-Q1-K1-A8

Method BLASTX
NCBI GI g445612
BLAST score 595
E value 2.0e-61
Match length 122
% identity 93

NCBI Description ribosomal protein S19 [Solanum tuberosum]

Seq. No. 6943

Contig ID 7624 1.R1040

5'-most EST LIB3039-054-Q1-E1-D4

Method BLASTX
NCBI GI g2583135
BLAST score 239
E value 6.0e-20
Match length 86
% identity 56

NCBI Description (AC002387) hypothetical protein [Arabidopsis thaliana]

>gi 3822216 (AF074948) FIL [Arabidopsis thaliana]

>gi_4322477_gb AAD16053 (AF087015) abnormal floral organs

protein [Arabidopsis thaliana]

Seq. No. 6944

Contig ID 7626_1.R1040 5'-most EST bnc700605624.h2

Seq. No. 6945

Contig ID 7627 1.R1040

5'-most EST fC-gmro7000746976f1

Method BLASTX
NCBI GI g2501433
BLAST score 542
E value 5.0e-55
Match length 172
% identity 59

NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-24 KD (UBIQUITIN-PROTEIN

LIGASE) (UBIQUITIN CARRIER PROTEIN) (E2-EPF5) >gi 181916

(M91670) ubiquitin carrier protein [Homo sapiens]

Seq. No. 6946

Contig ID 7627_2.R1040 5'-most EST leu701146523.h1

Method BLASTX
NCBI GI g345829
BLAST score 381
E value 1.0e-36
Match length 108
% identity 62

NCBI Description ubiquitin carrier protein E2 - human

Seq. No. 6947

Contig ID 7627 3.R1040

5'-most EST fC-gmro7000746976r1

Match length

% identity

130

57

```
Seq. No.
                    6948
Contig ID
                    7634 1.R1040
5'-most EST
                    epx701108179.h1
Method
                    BLASTX
NCBI GI
                    q4455293
BLAST score
                    211
E value
                   1.0e-16
Match length
                   57
% identity
                   72
NCBI Description
                   (AL035528) putative protein [Arabidopsis thaliana]
Seq. No.
                   6949
Contig ID
                   7636 1.R1040
5'-most EST
                   LIB3039-054-Q1-E1-B6
                   6950
Seq. No.
Contig ID
                   7637 1.R1040
5'-most EST
                   LIB3049-016-Q1-E1-F11
Seq. No.
                   6951
Contig ID
                   7637 2.R1040
5'-most EST
                   LIB3093-010-Q1-K1-C1
Seq. No.
                   6952
Contig ID
                   7638 1.R1040
5'-most EST
                   LIB3039-010-Q1-E1-D2
Seq. No.
                   6953
Contig ID
                   7638 2.R1040
5'-most EST
                   ncj7\overline{0}0981237.h1
Seq. No.
                   6954
Contig ID
                   7643 1.R1040
5'-most EST
                   LIB3106-073-Q1-K1-C6
Method
                   BLASTX
NCBI GI
                   g4263782
BLAST score
                   146
E value
                   5.0e-09
Match length
                   47
% identity
                   53
NCBI Description (AC006068) hypothetical protein [Arabidopsis thaliana]
                   6955
Seq. No.
Contig ID
                   7645 1.R1040
5'-most EST
                   jsh7\overline{0}1067491.h1
                   6956
Seq. No.
Contig ID
                   7647 1.R1040
5'-most EST
                   LIB3049-046-Q1-E1-F3
Method
                   BLASTX
NCBI GI
                   g3341443
BLAST score
                   376
E value
                   4.0e-36
```

1290

NCBI Description (AJ223074) acid phosphatase [Glycine max]

Seq. No.

6963

```
Seq. No.
                   6957
 Contig ID
                   7648 1.R1040
 5'-most EST
                   LIB3039-026-Q1-E1-D11
Method
                   BLASTX
NCBI GI
                   q4220524
BLAST score
                   291
E value
                   7.0e-26
Match length
                   148
% identity
                   47
NCBI Description
                  (AL035356) putative protein [Arabidopsis thaliana]
Seq. No.
                   6958
Contig ID
                   7648 2.R1040
5'-most EST
                   LIB3039-016-Q1-E1-H4
Seq. No.
                   6959
Contig ID
                   7652 1.R1040
5'-most EST
                   LIB3074-025-Q1-E1-E11
Method
                   BLASTN
NCBI GI
                   g2852444
BLAST score
                   119
E value
                   6.0e-60
Match length
                   319
% identity
                   84
NCBI Description Salix bakko mRNA for SUI1 homolog, complete cds
Seq. No.
                   6960
Contig ID
                   7652 2.R1040
5'-most EST
                  LIB3106-061-Q1-K1-C2
Seq. No.
                   6961
Contig ID
                   7657 1.R1040
5'-most EST
                  g4296782
Method
                  BLASTX
NCBI GI
                  g3947719
BLAST score
                   237
E value
                   9.0e-20
Match length
                  51
% identity
                  92
NCBI Description
                  (AJ012653) ribosomal protein S28 [Prunus persica]
                  >gi_3947721_emb_CAA10102_ (AJ012654) ribosomal protein S28
                   [Prunus persica] >gi_3947723_emb CAA10103 (AJ012655)
                  ribosomal protein S28 [Prunus persica]
Seq. No.
                  6962
Contig ID
                  7657 2.R1040
5'-most EST
                  LIB3039-001-Q1-E1-D7
Method
                  BLASTN
NCBI GI
                  g3947720
BLAST score
                  66
E value
                  1.0e-28
Match length
                  182
% identity
                  84
NCBI Description Prunus persica mRNA for ribosomal protein S28 (rps28.2)
```

```
Contig ID
                   7657 3.R1040
5'-most EST
                   jsh701070119.h1
Method
                   BLASTN
NCBI GI
                   q3947720
BLAST score
                   66
E value
                   1.0e-28
Match length
                   182
% identity
                   84
NCBI Description Prunus persica mRNA for ribosomal protein S28 (rps28.2)
Seq. No.
                   6964
Contig ID
                   7657 4.R1040
5'-most EST
                   LIB3049-049-Q1-E1-E3
Method
                   BLASTX
NCBI GI
                   g3947719
BLAST score
                   237
E value
                   1.0e-19
Match length
                   51
% identity
                   92
NCBI Description
                   (AJ012653) ribosomal protein S28 [Prunus persica]
                   >gi 3947721 emb CAA10102 (AJ012654) ribosomal protein S28
                   [Prunus persica] >gi_3947723 emb_CAA10103_ (AJ012655)
                   ribosomal protein S28 [Prunus persica]
Seq. No.
                   6965
Contig ID
                   7657 7.R1040
                   gsv7\overline{0}1050725.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3947719
BLAST score
                   229
E value
                   7.0e-19
Match length
                   51
% identity
NCBI Description
                   (AJ012653) ribosomal protein S28 [Prunus persica]
                   >gi_3947721 emb CAA10102 (AJ012654) ribosomal protein S28
                   [Prunus persica] >gi_3947723 emb CAA10103 (AJ012655)
                   ribosomal protein S28 [Prunus persica]
Seq. No.
                   6966
Contig ID
                  7659 1.R1040
5'-most EST
                  LIB3170-051-Q1-K2-F2
Method
                  BLASTX
NCBI GI
                  g4510379
BLAST score
                  387
E value
                  5.0e-37
Match length
                  182
% identity
                  48
NCBI Description
                  (AC007017) hypothetical protein [Arabidopsis thaliana]
```

Seq. No. 6967

Contig ID 7665_1.R1040

5'-most EST LIB3170-034-Q1-K1-H12

Method BLASTN
NCBI GI g19572
BLAST score 150
E value 2.0e-78
Match length 230

% identity

NCBI Description M.sativa C29 mRNA for snRNP-related protein

Seq. No. 6968

Contig ID 7665 2.R1040

5'-most EST LIB3072-007-Q1-E1-G1

Method BLASTN NCBI GI g19572 BLAST score 137 E value 5.0e-71 Match length 229 % identity 90

NCBI Description M.sativa C29 mRNA for snRNP-related protein

Seq. No. 6969

Contig ID 7666 1.R1040

5'-most EST LIB3073-015-Q1-K1-B9

Method BLASTN NCBI GI g313026 BLAST score 86 E value 2.0e-40 Match length 210 % identity 85

NCBI Description L.esculentum rpl38 mRNA for ribosomal protein L38

Seq. No. 6970

Contig ID 7667 1.R1040 5'-most EST gsv701054182.h1

Method BLASTX NCBI GI q1587206 BLAST score 961 E value 1.0e-104 Match length 203 % identity 91

NCBI Description T complex protein [Cucumis sativus]

Seq. No. 6971

Contig ID 7667 2.R1040

5'-most EST LIB3138-126-Q1-N1-C12

Method BLASTX NCBI GI q3024697 BLAST score 1420 E value 1.0e-158 Match length 325 % identity 84

NCBI Description T-COMPLEX PROTEIN 1, EPSILON SUBUNIT (TCP-1-EPSILON)

(CCT-EPSILON) >gi_2213618 (AC000103) F21J9.12 [Arabidopsis

thaliana]

Seq. No. 6972

Contig ID 7667 3.R1040 5'-most EST pcp700992489.h1

Method BLASTX NCBI GI g1587206 BLAST score 530 E value 4.0e-54 Match length 127



67

6974

```
% identity
NCBI Description T complex protein [Cucumis sativus]
Seq. No.
                   6973
Contig ID
                   7668 1.R1040
5'-most EST
                  LIB3138-028-01-N1-E1
Method
                  BLASTX
NCBI GI
                  q4469025
BLAST score
                   488
E value
                   4.0e-49
Match length
                  106
```

NCBI Description (AL035602) putative protein [Arabidopsis thaliana]

Contig ID 7668 2.R1040 5'-most EST leu701147156.h1 Method BLASTX NCBI GI q4469025 BLAST score 408 E value 8.0e-40 98

Match length % identity 60

% identity

Seq. No.

NCBI Description (AL035602) putative protein [Arabidopsis thaliana]

Seq. No. 6975

Contig ID 7668 3.R1040 5'-most EST dpv701101316.hl

Method BLASTX NCBI GI q4469025 BLAST score 200 2.0e-15 E value Match length 43 % identity 84

NCBI Description (AL035602) putative protein [Arabidopsis thaliana]

Seq. No. 6976

Contig ID 7669 1.R1040

5'-most EST uC-gmropic091q08b1

Method BLASTX NCBI GI g730463 BLAST score 319 E value 4.0e-29 Match length 105 % identity 58

NCBI Description 60S RIBOSOMAL PROTEIN L37B (YL37) (RP47)

>gi_630323_pir__S44069 ribosomal protein L35a.e.c15 - yeast (Saccharomyces cerevisiae) >gi_484241 (L23923) ribosomal

protein L37 [Saccharomyces cerevisiae]

>gi 1420537 emb CAA99454 (Z75142) ORF YOR234c

[Saccharomyces cerevisiae]

Seq. No. 6977

Contig ID 7669 2.R1040

5'-most EST LIB3040-052-Q1-E1-B9

Method BLASTX NCBI GI g730463



BLAST score 319 E value 3.0e-29 Match length 105 % identity 57

NCBI Description 60S RIBOSOMAL PROTEIN L37B (YL37) (RP47)

>gi_630323_pir__S44069 ribosomal protein L35a.e.c15 - yeast
(Saccharomyces cerevisiae) >gi 484241 (L23923) ribosomal

protein L37 [Saccharomyces cerevisiae]

>gi_1420537_emb_CAA99454_ (Z75142) ORF YOR234c

[Saccharomyces cerevisiae]

Seq. No. 6978

Contig ID 7672 1.R1040

5'-most EST LIB3040-048-Q1-E1-D2

Method BLASTN
NCBI GI g643073
BLAST score 222
E value 1.0e-121
Match length 430
% identity 88

NCBI Description Fragaria x ananassa putative 40S ribosomal protein s12

mRNA, complete cds

Seq. No. 6979

Contig ID 7672 2.R1040

5'-most EST LIB3055-004-Q1-N1-H6

Method BLASTN
NCBI GI g643073
BLAST score 188
E value 1.0e-101
Match length 428
% identity 86

NCBI Description Fragaria x ananassa putative 40S ribosomal protein s12

mRNA, complete cds

Seq. No. 6980

Contig ID 7672_4.R1040 5'-most EST sat701013854.h1

Method BLASTX
NCBI GI g1173187
BLAST score 279
E value 4.0e-44
Match length 139
% identity 75

NCBI Description 40S RIBOSOMAL PROTEIN S23 (S12) >gi_1362041_pir__S56673

ribosomal protein S23.e, cytosolic (clone RJ3) - garden strawberry >gi_643074 (U19940) putative 40S ribosomal

protein s12 [Fragaria x ananassa]

Seq. No. 6981

Contig ID 7672_5.R1040 5'-most EST sat701010376.h1

Method BLASTN
NCBI GI g643073
BLAST score 73
E value 9.0e-33
Match length 141



% identity 88

NCBI Description Fragaria x ananassa putative 40S ribosomal protein s12

mRNA, complete cds

Seq. No. 6982

Contig ID 7677_1.R1040

5'-most EST jC-gmf102220068c02a1

Method BLASTX
NCBI GI g4335750
BLAST score 252
E value 2.0e-21
Match length 76
% identity 54

NCBI Description (AC006284) putative beta-1,3-endoglucanase [Arabidopsis

thaliana]

Seq. No. 6983

Contig ID 7677_2.R1040 5'-most EST sat701009283.h1

Seq. No. 6984

Contig ID 7678 1.R1040

5'-most EST LIB3039-053-Q1-E1-E9

Seq. No. 6985

Contig ID 7681_1.R1040 5'-most EST gsv701054724.h1

Method BLASTX
NCBI GI g4406372
BLAST score 1532
E value 1.0e-171
Match length 378
% identity 76

NCBI Description (AF109156) thiosulfate sulfurtransferase [Datisca

glomerata]

Seq. No. 6986

Contig ID 7683_1.R1040

5'-most EST uC-gmropic046a05b1
Method BLASTX

Method BLASTX
NCBI GI g4558661
BLAST score 1198
E value 1.0e-132
Match length 294
% identity 78

NCBI Description (AC007063) putative malate oxidoreductase (NAD)

[Arabidopsis thaliana]

Seq. No. 6987

Contig ID 7684_1.R1040

5'-most EST LIB3049-004-Q1-E1-G7

Method BLASTX
NCBI GI g3668082
BLAST score 619
E value 4.0e-64
Match length 204
% identity 62



NCBI Description (AC004667) putative DAL1 protein [Arabidopsis thaliana]

Seq. No. 6988

Contig ID 7684_2.R1040

5'-most EST LIB3039-033-Q1-E1-D9

Method BLASTX
NCBI GI g2459425
BLAST score 300
E value 4.0e-27
Match length 90
% identity 67

NCBI Description (AC002332) plastid protein [Arabidopsis thaliana]

Seq. No. 6989

Contig ID 7690 1.R1040

5'-most EST LIB3073-023-Q1-K1-G9

Method BLASTX
NCBI GI g294060
BLAST score 192
E value 3.0e-14
Match length 127
% identity 33

NCBI Description (L06467) major latex protein [Papaver somniferum]

Seq. No. 6990

Contig ID 7691 1.R1040

5'-most EST LIB3039-053-Q1-E1-E8

Method BLASTX
NCBI GI g99755
BLAST score 341
E value 4.0e-32
Match length 128
% identity 53

NCBI Description RNA-directed DNA polymerase (EC 2.7.7.49) - Arabidopsis

thaliana retrotransposon Tal-1 (fragment)

>gi_16356_emb_CAA37917_ (X53973) reverse transcriptase

[Arabidopsis thaliana]

Seq. No. 6991

Contig ID 7692_1.R1040

5'-most EST LIB3039-053-Q1-E1-C1

Method BLASTX
NCBI GI g4185140
BLAST score 410
E value 4.0e-40
Match length 88
% identity 92

NCBI Description (AC005724) putative small nuclear ribonucleoprotein E

(snRNP-E) [Arabidopsis thaliana]

Seq. No. 6992

Contig ID 7692_2.R1040

5'-most EST LIB3072-008-Q1-E1-E1

Method BLASTX
NCBI GI g4185140
BLAST score 409
E value 4.0e-40

Match length 88 % identity 91

NCBI Description (AC005724) putative small nuclear ribonucleoprotein E

(snRNP-E) [Arabidopsis thaliana]

Seq. No. 6993

Contig ID 7692 3.R1040

5'-most EST LIB3170-080-Q1-K1-A12

Method BLASTX
NCBI GI g4185140
BLAST score 378
E value 2.0e-36
Match length 81
% identity 93

NCBI Description (AC005724) putative small nuclear ribonucleoprotein E

(snRNP-E) [Arabidopsis thaliana]

Seq. No. 6994

Contig ID 7696_1.R1040

5'-most EST LIB3056-003-Q1-N1-G2

Method BLASTX
NCBI GI g1730035
BLAST score 231
E value 7.0e-19
Match length 121
% identity 40

NCBI Description PROTEIN KE2 >gi_2137410_pir__I53651 hydrophilic protein -

mouse >gi_198574 (M65255) hydrophilic protein [Mus musculus] >gi_198576 (M65256) hydrophilic protein [Mus musculus] >gi_3811379 (AF100956) KE2 [Mus musculus]

>gi_4050102 (AF110520) KE2 [Mus musculus]

Seq. No. 6995

Contig ID 7696 2.R1040 5'-most EST 6HC-01-Q1-B1-B12

Method BLASTX
NCBI GI g2344898
BLAST score 571
E value 1.0e-58
Match length 168
% identity 65

NCBI Description (AC002388) 60S ribosomal protein L30 isolog [Arabidopsis

thaliana]

Seq. No. 6996

Contig ID 7698 1.R1040

5'-most EST LIB3039-053-Q1-E1-C8

Seq. No. 6997

Contig ID 7699_1.R1040

5'-most EST LIB3040-018-Q1-E1-G4

Method BLASTX
NCBI GI g2431771
BLAST score 248
E value 7.0e-21
Match length 114
% identity 49

```
NCBI Description
                   (U62753) acidic ribosomal protein P2b [Zea mays]
 Seq. No.
                    6998
 Contig ID
                    7699 2.R1040
 5'-most EST
                   LIB3106-077-P1-K1-A7
 Method
                   BLASTX
 NCBI GI
                   g2431771
 BLAST score
                   226
 E value
                   2.0e-18
 Match length
                   114
 % identity
                   45
 NCBI Description
                   (U62753) acidic ribosomal protein P2b [Zea mays]
 Seq. No.
                   6999
 Contig ID
                   7699 3.R1040
 5'-most EST
                   gsv701048437.hl
Method
                   BLASTX
 NCBI GI
                   g2431771
BLAST score
                   162
E value
                   2.0e-11
Match length
                   47
 % identity
                   68
NCBI Description
                   (U62753) acidic ribosomal protein P2b [Zea mays]
Seq. No.
                   7000
                   7699_4.R1040
Contig ID
5'-most EST
                   wrg700792224.h1
Method
                   BLASTX
NCBI GI
                   g2431771
BLAST score
                   251
                   1.0e-21
E value
Match length
                   70
% identity
                   70
NCBI Description
                   (U62753) acidic ribosomal protein P2b [Zea mays]
Seq. No.
                   7001
Contig ID
                   7700 1.R1040
5'-most EST
                   LIB3109-002-Q1-K1-B11
Method
                   BLASTX
NCBI GI
                   q3175990
BLAST score
                   2186
E value
                   0.0e+00
Match length
                   436
% identity
                   94
NCBI Description
                   (AJ005836) GDP dissociation inhibitor [Cicer arietinum]
Seq. No.
                   7002
Contig ID
                   7700 2.R1040
5'-most EST
                  wvk700685939.h1
Method
                  BLASTN
NCBI GI
                  g3175989
BLAST score
                  247
E value
                  1.0e-136
Match length
                  430
% identity
                  90
NCBI Description Cicer arietinum mRNA for GDP dissociation inhibitor (gdi)
```

% identity

NCBI Description

81



```
Seq. No.
                   7003
Contig ID
                   7700 3.R1040
                   LIB3049-032-Q1-E1-A9
5'-most EST
Method
                   BLASTN
NCBI GI
                   q3175989
BLAST score
                   147
E value
                   7.0e-77
Match length
                   294
% identity
NCBI Description Cicer arietinum mRNA for GDP dissociation inhibitor (gdi)
Seq. No.
                   7004
                   7704 1.R1040
Contig ID
5'-most EST
                   LIB3073-024-Q1-K1-H9
Seq. No.
                   7005
Contig ID
                   7706 1.R1040
                   LIB3065-001-Q1-N1-G12
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4263521
BLAST score
                   351
E value
                   3.0e-33
Match length
                   87
% identity
                   15
                   (AC004044) putative WD-repeat protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7006
Contig ID
                   7708 1.R1040
5'-most EST
                   LIB3039-053-Q1-E1-B3
Seq. No.
                   7007
                   7713 1.R1040
Contig ID
5'-most EST
                   LIB3039-041-Q1-E1-C12
Seq. No.
                   7008
Contig ID
                   7715 1.R1040
5'-most EST
                   uC-gmflminsoy077b08b1
Method
                   BLASTX
NCBI GI
                   g3785986
BLAST score
                   154
E value
                   7.0e-10
Match length
                   65
% identity
                   51
NCBI Description
                   (AC005560) RGA1 protein [Arabidopsis thaliana]
                   7009
Seq. No.
Contig ID
                   7727 1.R1040
5'-most EST
                   fC-gmse700753078a2
Method
                   BLASTX
NCBI GI
                   g3702326
BLAST score
                   839
E value
                  7.0e-90
Match length
                  194
```

1300

(AC005397) hypothetical protein [Arabidopsis thaliana]



```
7010
Seq. No.
Contig ID
                  7727 2.R1040
5'-most EST
                  gsv701055260.hl
Method
                  BLASTX
NCBI GI
                  g3702326
BLAST score
                  237
                  2.0e-26
E value
Match length
                  104
% identity
                   62
NCBI Description
                  (AC005397) hypothetical protein [Arabidopsis thaliana]
                  7011
Seq. No.
Contig ID
                  7729 1.R1040
5'-most EST
                  LIB3040-021-Q1-E1-D10
Method
                  BLASTX
NCBI GI
                  g3320379
BLAST score
                  924
E value
                  1.0e-100
Match length
                  204
% identity
                  88
NCBI Description
                  (AF014375) putative JUN kinase activation domain binding
                  protein [Medicago sativa]
                  7012
Seq. No.
                  7733 1.R1040
Contig ID
5'-most EST
                  LIB3039-052-Q1-E1-G12
Method
                  BLASTX
NCBI GI
                  g2583108
BLAST score
                  651
                  9.0e-68
E value
                  308
Match length
% identity
                  44
NCBI Description
                  (AC002387) putative surface protein [Arabidopsis thaliana]
Seq. No.
                  7013
Contig ID
                  7737 1.R1040
5'-most EST
                  LIB3107-056-Q1-K1-G7
                  7014
Seq. No.
Contig ID
                  7746 1.R1040
5'-most EST
                  LIB3049-001-Q1-E1-F6
Method
                  BLASTX
NCBI GI
                  g3687237
BLAST score
                  885
E value
                  1.0e-107
Match length
                  354
                  59
% identity
NCBI Description
                  (AC005169) putative Cys3His zinc-finger protein
                   [Arabidopsis thaliana]
```

Seq. No. 70

Seq. No. 7015 Contig ID 7746 2.R1040

5'-most EST LIB3049-048-01-E1-C7

Method BLASTX
NCBI GI g3687237
BLAST score 383
E value 1.0e-36



```
Match length
% identity
                  (AC005169) putative Cys3His zinc-finger protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                  7016
                  7748 1.R1040
Contig ID
                  LIB3039-052-Q1-E1-C1
5'-most EST
```

Seq. No. 7757 1.R1040 Contig ID seb700652631.hl 5'-most EST BLASTX Method q4335725 NCBI GI 280 BLAST score 2.0e-24

7017

E value Match length 142 46 % identity

NCBI Description (AC006248) hypothetical protein [Arabidopsis thaliana]

7018 Seq. No. 7757 2.R1040 Contig ID jC-gmle01810001h12d1 5'-most EST

7019 Seq. No. 7758 1.R1040 Contig ID LIB3039-052-Q1-E1-C9 5'-most EST BLASTX Method q4263704 NCBI GI BLAST score 280

9.0e-25 E value 103 Match length % identity 48

(AC006223) putative sugar starvation-induced protein NCBI Description [Arabidopsis thaliana]

7020 Seq. No. 7774 1.R1040 Contig ID LIB3039-051-Q1-E1-G10 5'-most EST BLASTN Method g170091 NCBI GI

183 BLAST score 1.0e-98 E value 287 Match length 92 % identity

NCBI Description Glycine max vegetative storage protein (vspB) gene,

complete cds

7021 Seq. No. 7776 1.R1040 Contig ID

LIB3049-047-Q1-E1-G5 5'-most EST

BLASTN Method g1531753 NCBI GI 140 BLAST score 2.0e-72 E value 400 Match length % identity 84

1302



NCBI Description A.officinalis mRNA for Histone H3

Seq. No. 7022

Contig ID 7776 2.R1040

5'-most EST LIB3170-020-Q1-J1-C12

Method BLASTX
NCBI GI g19611
BLAST score 467
E value 1.0e-46
Match length 96
% identity 99

NCBI Description (X13675) histone H3 (AA 1-123) [Medicago sativa]

>gi 2916748 emb_CAA05554_ (AJ002555) histone H3 [Pisum

sativum]

Seq. No. 7023

Contig ID 7776_3.R1040

5'-most EST LIB3170-015-Q1-J1-E5

Seq. No. 7024

Contig ID 7783_1.R1040

5'-most EST LIB3039-047-Q1-E1-B8

Seq. No. 7025

Contig ID 7796_1.R1040 5'-most EST wrg700791377.h1

Method BLASTX
NCBI GI g3885513
BLAST score 425
E value 1.0e-41
Match length 88
% identity 88

NCBI Description (AF084201) similar to chloroplast 50S ribosomal protein L31

[Medicago sativa]

Seq. No. 7026

Contig ID 7803_1.R1040 5'-most EST epx701109040.h1

Seq. No. 7027

Contig ID 7806_1.R1040

5'-most EST LIB3039-051-Q1-E1-D4

Method BLASTX
NCBI GI 94226129
BLAST score 175
E value 2.0e-12
Match length 155
% identity 30

NCBI Description (AF125459) No definition line found [Caenorhabditis

elegans]

Seq. No. 7028

Contig ID 7807 1.R1040

5'-most EST LIB3039-051-Q1-E1-C10

Seq. No. 7029

Contig ID 7808 1.R1040

1303



5'-most EST sat701006613.h1

Seq. No. 7030

Contig ID 7809 1.R1040 5'-most EST leu701150109.h1

Seq. No. 7031

Contig ID 7811 1.R1040 5'-most EST leu701146266.h1

Seq. No. 7032

Contig ID 7811_2.R1040 5'-most EST ncj700978210.h1

Seq. No. 7033

Contig ID 7812_1.R1040

5'-most EST LIB3039-020-Q1-E1-B8

Method BLASTX
NCBI GI g1708464
BLAST score 257
E value 3.0e-22
Match length 90

% identity 57

NCBI Description PUTATIVE DIHYDROXY-ACID DEHYDRATASE PRECURSOR (DAD)

(2,3-DIHYDROXY ACID HYDROLYASE) >gi_1213255_emb_CAA93689_

(Z69795) unknown [Schizosaccharomyces pombe]

Seq. No. 7034

Contig ID 7814 1.R1040 5'-most EST leu701144777.h1

Method BLASTX
NCBI GI 9400649
BLAST score 145
E value 8.0e-09
Match length 65
% identity 40

NCBI Description NADH-UBIQUINONE OXIDOREDUCTASE 12 KD SUBUNIT PRECURSOR

(COMPLEX I-12KD) (CI-12KD) >gi_479205_pir__S32568 gene NUO-12.3 protein - Neurospora crassa >gi_3040_emb_CAA48768_

(X68965) NUO-12.3 [Neurospora crassa]

Seq. No. 7035

Contig ID 7815_1.R1040 5'-most EST uxk700670174.h1

Seq. No. 7036

Contig ID 7820 1.R1040

5'-most EST LIB3039-051-Q1-E1-B9

Method BLASTX
NCBI GI g2827888
BLAST score 594
E value 2.0e-61
Match length 197
% identity 62

NCBI Description (AF016621) ATP-dependent Clp protease proteolytic subunit

[Arabidopsis thaliana]



```
Seq. No.
                   7037
                   7829 1.R1040
Contig ID
5'-most EST
                   kl1701205406.h1
                   7038
Seq. No.
                   7837 1.R1040
Contig ID
5'-most EST
                   zhf700964782.h1
                   7039
Seq. No.
                   7841 1.R1040
Contig ID
                   LIB3039-050-Q1-E1-G12
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3821254
BLAST score
                   1529
                   0.0e+00
E value
                   463
Match length
% identity
NCBI Description (AJ007789) geranylgeranyl reductase [Nicotiana tabacum]
Seq. No.
                   7846 1.R1040
Contig ID
5'-most EST
                   LIB3039-050-Q1-E1-E2
                   7041
Seq. No.
Contig ID
                   7852 1.R1040
5'-most EST
                   uC-gmflminsoy057c11b1
Seq. No.
                   7042
Contig ID
                   7855 1.R1040
5'-most EST
                   LIB3039-035-Q1-E1-G1
Method
                   BLASTX
NCBI GI
                   g464621
BLAST score
                   819
E value
                   1.0e-87
Match length
                   224
% identity
                   71
NCBI Description 60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi 280374 pir__S28586
                   ribosomal protein ML16 - common ice plant >gi_19539_emb_CAA49175_ (X69378) ribosomal protein YL16
                   [Mesembryanthemum crystallinum]
Seq. No.
                   7043
Contig ID
                   7855 2.R1040
5'-most EST
                   g5676862
Method
                   BLASTX
NCBI GI
                   g464621
```

BLAST score 323 E value 7.0e-58 Match length 199

% identity 78

NCBI Description 60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi 280374 pir __S28586

ribosomal protein ML16 - common ice plant

>gi 19539 emb CAA49175 (X69378) ribosomal protein YL16

[Mesembryanthemum crystallinum]

7044 Seq. No.

7855 4.R1040 Contig ID



```
5'-most EST
                   LIB3040-060-Q1-E1-G2
Method
                   BLASTX
NCBI GI
                   q464621
BLAST score
                   233
E value
                   2.0e-19
Match length
                   56
                   80
% identity
NCBI Description
                   60S RIBOSOMAL PROTEIN L6 (YL16-LIKE) >gi_280374_pir__S28586
                   ribosomal protein ML16 - common ice plant
>gi_19539_emb_CAA49175_ (X69378) ribosomal protein YL16
                   [Mesembryanthemum crystallinum]
Seq. No.
                   7045
                   7856 1.R1040
Contig ID
                   LIB3106-058-Q1-K1-F4
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3097266
BLAST score
                   1042
E value
                   0.0e + 00
                   407
Match length
                   88
% identity
NCBI Description (Y17186) translation initiation factor [Pisum sativum]
                   7046
Seq. No.
Contig ID
                   7858 1.R1040
5'-most EST
                   LIB3039-050-Q1-E1-D10
Method
                   BLASTX
NCBI GI
                   g3036807
BLAST score
                   664
E value
                   4.0e-85
Match length
                   269
                   62
% identity
NCBI Description (AL022373) putative protein [Arabidopsis thaliana]
Seq. No.
                   7047
Contig ID
                   7860 1.R1040
5'-most EST
                   LIB3039-050-Q1-E1-D12
                   BLASTN
Method
NCBI GI
                   q3980263
BLAST score
                   355
                   0.0e + 00
E value
                   814
Match length
                   89
% identity
NCBI Description Cicer arietinum mRNA for 20S proteasome beta subunit
                   7048
Seq. No.
                   7860 2.R1040
Contig ID
5'-most EST
                   seb700652801.h1
Method
                   BLASTN
                   g3980263
NCBI GI
BLAST score
                   136
E value
                   3.0e-70
```

Match length 456 % identity 89

NCBI Description Cicer arietinum mRNA for 20S proteasome beta subunit

7049 Seq. No.



```
Contig ID
                  7868 1.R1040
5'-most EST
                  LIB3093-035-Q1-K1-A8
                  BLASTX
Method
NCBI GI
                  g3063454
BLAST score
                  148
                  4.0e-09
E value
Match length
                  136
                  46
% identity
NCBI Description (AC003981) F22013.16 [Arabidopsis thaliana]
Seq. No.
                  7050
                  7868 2.R1040
Contig ID
5'-most EST
                  kmv700739710.h1
                  7051
Seq. No.
                  7875 1.R1040
Contig ID
5'-most EST
                  awf700839764.h1
                  BLASTX
Method
NCBI GI
                  g4567302
BLAST score
                  148
                  4.0e-09
E value
Match length
                  86
% identity
NCBI Description (AC005956) unknown protein [Arabidopsis thaliana]
                  7052
Seq. No.
Contig ID
                  7877 1.R1040
5'-most EST
                  gsv701047796.h1
Method
                  BLASTX
NCBI GI
                  g2497543
BLAST score
                  2263
E value
                  0.0e+00
Match length
                  493
% identity
                  89
NCBI Description PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi 542061 pir S41379
                  pyruvate kinase - common tobacco >gi 444023 emb CAA82628
                   (Z29492) pyruvate kinase [Nicotiana tabacum]
Seq. No.
                  7053
                  7877 2.R1040
Contig ID
5'-most EST
                   jC-gmst02400055d10a1
Method
                  BLASTX
NCBI GI
                  g4033431
BLAST score
                  570
                  8.0e-59
E value
```

126 Match length

NCBI Description PROBABLE PYRUVATE KINASE, CYTOSOLIC ISOZYME (PK)

>gi 2982467 emb CAA18231 (AL022223) pyruvate kinase like

protein [Arabidopsis thaliana]

7054 Seq. No.

% identity

7877 3.R1040 Contig ID

LIB3170-061-Q1-K2-D11 5'-most EST

Method BLASTX NCBI GI g2497543 BLAST score 495



E value 5.0e-50 Match length 104 % identity 89

NCBI Description PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi_542061_pir__S41379 pyruvate kinase - common tobacco >gi_444023_emb_CAA82628_

(Z29492) pyruvate kinase [Nicotiana tabacum]

Seq. No. 7055

Contig ID 7877_4.R1040 5'-most EST sat701010766.h1

Method BLASTX
NCBI GI g2497543
BLAST score 436
E value 6.0e-43
Match length 143
% identity 75

NCBI Description PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi_542061_pir__S41379

pyruvate kinase - common tobacco >gi_444023_emb_CAA82628_

(Z29492) pyruvate kinase [Nicotiana tabacum]

Seq. No. 7056

Contig ID 7880_1.R1040 5'-most EST leu701146874.h1

Seq. No. 7057

Contig ID 7880_2.R1040

5'-most EST LIB3170-012-Q1-J1-C11

Seq. No. 7058

Contig ID 7883_1.R1040 5'-most EST taw700655766.h1

Seq. No. 7059

Contig ID 7883_2.R1040 5'-most EST g5510176

Seq. No.

7060

Contig ID 7883 3.R1040

5'-most EST LIB3073-015-Q1-K1-B1

Seq. No. 7061

Contig ID 7883_4.R1040 5'-most EST ssr700557657.h1

Seq. No. 7062

Contig ID 7884 1.R1040

5'-most EST LIB3039-049-Q1-E1-H9

Method BLASTX
NCBI GI g4204294
BLAST score 202
E value 7.0e-16
Match length 71
% identity 66

NCBI Description (AC003027) lcl prt seq No definition line found

[Arabidopsis thaliana]

Seq. No. 7063



7892 1.R1040 Contig ID LIB3109-048-Q1-K1-A1 5'-most EST BLASTX Method q3063449 NCBI GI BLAST score 1364 1.0e-151 E value 347 Match length 74 % identity NCBI Description (AC003981) F22013.11 [Arabidopsis thaliana] 7064 Seq. No. 7892 2.R1040 Contig ID LIB3039-049-Q1-E1-H10 5'-most EST BLASTX Method g3063449 NCBI GI BLAST score 436 E value 3.0e-43 109 Match length % identity (AC003981) F22013.11 [Arabidopsis thaliana] NCBI Description 7065 Seq. No. 7896 1.R1040 Contig ID jC-gmst02400015c11a1 5'-most EST BLASTX Method q3668097 NCBI GI BLAST score 649 7.0e-68 E value 155 Match length % identity (AC004667) putative glycine cleavage system protein H NCBI Description precursor [Arabidopsis thaliana] 7066 Seq. No. 7896 2.R1040 Contig ID LIB3170-035-Q1-K1-A1 5'-most EST Method BLASTX q3668097 NCBI GI BLAST score 571 7.0e-59 E value Match length 138 % identity 78 (AC004667) putative glycine cleavage system protein H NCBI Description precursor [Arabidopsis thaliana] 7067 Seq. No. 7900 1.R1040 Contig ID LIB3106-010-Q1-K1-G1 5'-most EST BLASTX Method g1717755 NCBI GI 829 BLAST score

1.0e-88 E value

263 Match length 60 % identity NCBI Description

TROPINONE REDUCTASE HOMOLOG (P29X) >gi 539028 pir__C48674

tropinone reductase homolog - jimsonweed >gi_424158 (L20475) 29kDa protein; high homology to aa sequence of



tropinone reductases [Datura stramonium]

7068 Seq. No. 7900 3.R1040 Contig ID LIB3049-008-Q1-E1-F4 5'-most EST BLASTX Method NCBI GI g1717755 279 BLAST score 5.0e-27 E value 100 Match length 64 % identity TROPINONE REDUCTASE HOMOLOG (P29X) >gi_539028_pir C48674 NCBI Description tropinone reductase homolog - jimsonweed >gi_424158 (L20475) 29kDa protein; high homology to aa sequence of tropinone reductases [Datura stramonium] 7069 Seq. No. 7904 1.R1040 Contig ID LIB3049-020-Q1-E1-A8 5'-most EST BLASTX Method g1922938 NCBI GI 152 BLAST score 7.0e-10 E value 66 Match length 52 % identity (AC000106) Similar to Caenorhabditis hypothetical protein NCBI Description CO7A9.11 (gb_Z29094). [Arabidopsis thaliana] 7070 Seq. No. 7908 1.R1040 Contig ID LIB3039-049-Q1-E1-E4 5'-most EST BLASTX Method q1532163 NCBI GI 191 BLAST score 2.0e-29 E value 94 Match length % identity (U63815) similar to glutaredoxin encoded by GenBank NCBI Description Accession Number Z49699; localized according to blastn similarity to EST sequences; therefore, the coding span corresponds only to an area of similarity since the initation codon and stop c 7071 Seq. No. 7910 1.R1040 Contiq ID

leu701155340.h1 5'-most EST

BLASTX Method g3914999 NCBI GI 420 BLAST score 3.0e-41E value 110 Match length 71 % identity

SUPEROXIDE DISMUTASE, CHLOROPLAST [CU-ZN] PRECURSOR NCBI Description

>gi_1944326_dbj_BAA19675_ (D49486) copper/zinc-superoxide

dismutase precursor [Solidago canadensis]

7072 Seq. No.



7912 1.R1040 Contig ID gsv701046882.h1 5'-most EST BLASTX Method g4337176 NCBI GI 198 BLAST score 3.0e-15 E value Match length 136 13 % identity NCBI Description (AC006416) T31J12.4 [Arabidopsis thaliana] 7073 Seq. No. 7912 2.R1040 Contig ID LIB3039-049-Q1-E1-E8 5'-most EST BLASTX Method g4337176 NCBI GI 182 BLAST score 2.0e-13 E value 133 Match length 16 % identity (AC006416) T31J12.4 [Arabidopsis thaliana] NCBI Description 7074 Seq. No. 7916 1.R1040 Contig ID zpv700757882.h1 5'-most EST BLASTX Method g2245108 NCBI GI 530 BLAST score 7.0e-54 E value 188 Match length 53 % identity (Z97343) EREBP-4 homolog [Arabidopsis thaliana] NCBI Description Seq. No. 7075 7918 1.R1040 Contig ID hrw701060365.h15'-most EST BLASTX Method g3122703 NCBI GI BLAST score 624 6.0e-65 E value 151 Match length % identity 85 60S RIBOSOMAL PROTEIN L23A >gi 2641201 (AF031542) ribosomal NCBI Description protein L23a [Fritillaria agrestis] 7076 Seq. No. 7918 2.R1040 Contig ID LIB3094-047-Q1-K1-B3 5'-most EST 7077 Seq. No. 7919_1.R1040 Contig ID $LIB3\overline{0}39-049-Q1-E1-C8$ 5'-most EST BLASTX Method

Method BLASTX
NCBI GI g2330833
BLAST score 251
E value 3.0e-21
Match length 104
% identity 44



```
(Z98531) hypothetical protein [Schizosaccharomyces pombe]
NCBI Description
                  7078
Seq. No.
                  7919 2.R1040
Contig ID
5'-most EST
                  leu701155674.hl
                   7079
Seq. No.
                   7922 1.R1040
Contig ID
                  LIB3039-049-Q1-E1-D11
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1708971
                   381
BLAST score
                   2.0e-36
E value
                   160
Match length
% identity
                   (R)-MANDELONITRILE LYASE ISOFORM 1 PRECURSOR
NCBI Description
                   (HYDROXYNITRILE LYASE 1) ((R)-OXYNITRILASE 1)
                   >gi_421871_pir__S32156 mandelonitrile lyase (EC 4.1.2.10) -
                   black cherry >gi_288116_emb_CAA51194_ (X72617)
                   mandelonitrile lyase [Prunus serotina] >gi_1730332 (U78814)
                   (R)-(+)-mandelonitrile lyase isoform MDL1 precursor [Prunus
                   serotina] >gi 1090776_prf 2019441A mandelonitrile lyase
                   [Prunus serotina]
                   7080
Seq. No.
                   7934 1.R1040
Contig ID
5'-most EST
                   LIB3039-049-Q1-E1-C11
                   BLASTX
Method
                   q416731
NCBI GI
BLAST score
                   141
                   7.0e-09
E value
Match length
                   67
% identity
                   40
                   POLLEN SPECIFIC PROTEIN C13 PRECURSOR >gi 82655 pir JQ1107
NCBI Description
                   18.3K protein precursor, pollen - maize
                   >gi_255569_bbs_113677 (S44171) pollen specific protein [Zea
                   mays=corn, Peptide, 170 aa] [Zea mays]
                   >gi 1588669 prf 2209273A Zm13 [Zea mays]
                   7081
Seq. No.
                   7938 1.R1040
Contig ID
                   ncj7\overline{0}0983184.h1
5'-most EST
                   BLASTX
Method
                   g1653702
NCBI GI
                   275
BLAST score
                   3.0e-24
E value
                   81
Match length
 % identity
                   68
                   (D90915) dihydrolipoamide acetyltransferase component (E2)
NCBI Description
                   of pyruvate dehydrogenase complex [Synechocystis sp.]
 Seq. No.
                   7082
                   7938 2.R1040
 Contig ID
                   uC-qmropic073g06b1
 5'-most EST
```

BLASTX

211

g1653702

Method

NCBI GI BLAST score



E value 9.0e-17 Match length 50 % identity 82

NCBI Description (D90915) dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex [Synechocystis sp.]

Seq. No. 7083

Contig ID 7939 1.R1040

5'-most EST LIB3072-010-Q1-E1-A5

Method BLASTX
NCBI GI 94324409
BLAST score 233
E value 7.0e-19
Match length 206
% identity 31

NCBI Description (AF104119) intracellular chloride ion channel protein p64H1

[Rattus norvegicus]

Seq. No. 7084

Contig ID 7943 1.R1040

5'-most EST LIB3039-049-Q1-E1-B1

Method BLASTX
NCBI GI g3193284
BLAST score 390
E value 1.0e-37
Match length 154
% identity 55

NCBI Description (AF069298) No definition line found [Arabidopsis thaliana]

Seq. No. 7085

Contig ID 7944 1.R1040

5'-most EST LIB3039-049-Q1-E1-B10

Method BLASTN
NCBI GI 9703230
BLAST score 47
E value 2.0e-17
Match length 71
% identity 92

NCBI Description Pisum sativum (clone HW54) gene fragment

Seq. No. 7086

Contig ID 7947 1.R1040

5'-most EST LIB3039-048-Q1-E1-G12

Seq. No. 7087

Contig ID 7953 1.R1040

5'-most EST LIB3039-048-Q1-E1-G9

Seq. No. 7088

Contig ID 7955_1.R1040

5'-most EST LIB3049-008-Q1-E1-A7

Method BLASTX
NCBI GI g2252866
BLAST score 352
E value 1.0e-32
Match length 194
% identity 46



NCBI Description (AF013294) contains region of similarity to SYT [Arabidopsis thaliana]

Seq. No. 7089

Contig ID 7955 2.R1040 5'-most EST ncj700978551.h1

Method BLASTX
NCBI GI g2252866
BLAST score 266
E value 4.0e-23
Match length 56
% identity 95

NCBI Description (AF013294) contains region of similarity to SYT

[Arabidopsis thaliana]

Seq. No. 7090

Contig ID 7957 1.R1040

5'-most EST LIB3039-048-Q1-E1-F6

Method BLASTX
NCBI GI g4467111
BLAST score 173
E value 5.0e-12
Match length 63
% identity 60

NCBI Description (AL035538) putative protein [Arabidopsis thaliana]

Seq. No.

No. 7091

Contig ID 7961_1.R1040

5'-most EST LIB3039-048-Q1-E1-E5

Seq. No.

7092

Contig ID 7961_2.R1040 5'-most EST gsv701048205.h1

Seq. No.

7093

Contig ID 7964_1.R1040

5'-most EST uC-gmropic113a06b1

Method BLASTX
NCBI GI 9729704
BLAST score 339
E value 3.0e-31
Match length 223
% identity 12

NCBI Description DNA-BINDING PROTEIN HEXBP (HEXAMER-BINDING PROTEIN)

>gi_1078700_pir__A47156 hexamer-binding protein HEXBP -Leishmania major >gi_159342 (M94390) HEXBP DNA binding

protein [Leishmania major]

Seq. No. 7094

Contig ID 7975_1.R1040

5'-most EST LIB3039-048-Q1-E1-C6

Method BLASTX
NCBI GI g2341024
BLAST score 254
E value 5.0e-22
Match length 84
% identity 56



NCBI Description (AC000104) F19P19.1 [Arabidopsis thaliana]

Seq. No. 7095

Contig ID 7979_1.R1040

5'-most EST fC-gmle700557018f3

Method BLASTX
NCBI GI g1351271
BLAST score 1246
E value 1.0e-137
Match length 270
% identity 87

NCBI Description TRIOSEPHOSPHATE ISOMERASE CHLOROPLAST PRECURSOR (TIM)

>gi_1084309_pir__S52032 triose-phosphate isomerase (EC
5.3.1.1) precursor, chloroplast - spinach >gi_806312
(L36387) triosephosphate isomerase, chloroplast isozyme

[Spinacia oleracea]

Seq. No. 7096

Contig ID 7979 2.R1040

5'-most EST LIB3040-059-Q1-E1-H10

Method BLASTX
NCBI GI g1351271
BLAST score 701
E value 5.0e-74
Match length 182
% identity 76

NCBI Description TRIOSEPHOSPHATE ISOMERASE CHLOROPLAST PRECURSOR (TIM)

>gi_1084309_pir__S52032 triose-phosphate isomerase (EC
5.3.1.1) precursor, chloroplast - spinach >gi_806312
(L36387) triosephosphate isomerase, chloroplast isozyme

[Spinacia oleracea]

Seq. No. 7097

Contig ID 7979 3.R1040

5'-most EST fC-qmle7000739771d1

Method BLASTX
NCBI GI g1351271
BLAST score 347
E value 1.0e-32
Match length 78
% identity 87

NCBI Description TRIOSEPHOSPHATE ISOMERASE CHLOROPLAST PRECURSOR (TIM)

>gi_1084309_pir__S52032 triose-phosphate isomerase (EC
5.3.1.1) precursor, chloroplast - spinach >gi_806312
(L36387) triosephosphate isomerase, chloroplast isozyme

[Spinacia oleracea]

Seq. No. 7098

Contig ID 7979 4.R1040

5'-most EST LIB3107-068-Q1-K1-G9

Method BLASTX
NCBI GI g1174745
BLAST score 192
E value 1.0e-14
Match length 41
% identity 83

NCBI Description TRIOSEPHOSPHATE ISOMERASE, CHLOROPLAST PRECURSOR (TIM)

>gi 1363523 pir S53761 triose-phosphate isomerase (EC 5.3.1.1) precursor, chloroplast - rye >gi_609262_emb_CAA83533 (Z32521) triosephosphate isomerase [Secale cereale] >gi_1095494_prf__2109226B triosephosphate isomerase [Secale cereale]

Seq. No. Contig ID 7989 1.R1040 5'-most EST LIB3039-048-Q1-E1-B11 Method BLASTX NCBI GI q3080450 BLAST score 575 E value 3.0e-59

7099

Match length 187 % identity 60

NCBI Description (AL022605) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7100

7995 1.R1040 Contig ID

5'-most EST LIB3040-055-Q1-E1-E8

Seq. No.

7996 1.R1040 Contig ID

5'-most EST LIB3039-047-Q1-E1-G4

Seq. No. 7102

7997 1.R1040 Contig ID 5'-most EST leu701145158.h1

Method BLASTX NCBI GI q3687243 BLAST score 251 E value 3.0e-21 Match length 61 % identity 80

NCBI Description (AC005169) putative ribosomal protein [Arabidopsis

thaliana]

Seq. No. 7103

7997 2.R1040 Contig ID 5'-most EST sat701002975.hl

Method BLASTX g3687243 NCBI GI BLAST score 146 5.0e-09 E value Match length 32 % identity 91

NCBI Description (AC005169) putative ribosomal protein [Arabidopsis

thaliana]

Seq. No. 7104

Contig ID 8002 1.R1040

5'-most EST LIB3039-047-Q1-E1-H2

Method BLASTX NCBI GI g1071924 BLAST score 147 E value 1.0e-12 Match length 72

```
% identity
 NCBI Description
                   Kunitz trypsin inhibitor precursor - soybean
                   >gi_510515_emb_CAA56343_ (X80039) Kunitz trypsin inhibitor
                   [Glycine max]
 Seq. No.
                   7105
 Contig ID
                   8004 1.R1040
 5'-most EST
                   LIB3039-047-Q1-E1-H7
Method
                   BLASTX
NCBI GI
                   g3747050
BLAST score
                   491
E value
                   2.0e-49
Match length
                   118
 % identity
                   81
NCBI Description
                  (AF093540) ribosomal protein L26 [Zea mays]
Seq. No.
                   7106
Contig ID
                   8004 2.R1040
5'-most EST
                   LIB3170-077-Q1-J1-H5
Method
                   BLASTN
NCBI GI
                   g3747049
BLAST score
                   67
                   3.0e-29
E value
Match length
                   143
% identity
                   87
NCBI Description
                  Zea mays ribosomal protein L26 mRNA, partial cds
Seq. No.
                   7107
Contig ID
                   8006 1.R1040
5'-most EST
                   g5509188
Method
                   BLASTX
NCBI GI
                   g294845
BLAST score
                   637
E value
                   3.0e-66
Match length
                   262
% identity
                   49
NCBI Description
                   (L13655) membrane protein [Saccharum hybrid cultivar
                  H65-7052]
Seq. No.
                   7108
Contig ID
                   8006 2.R1040
5'-most EST
                  uC-gmropic058e08b1
Method
                  BLASTX
NCBI GI
                  g294845
BLAST score
                  248
E value
                  4.0e-21
Match length
                  87
% identity
                  59
NCBI Description
                  (L13655) membrane protein [Saccharum hybrid cultivar
                  H65-7052]
Seq. No.
                  7109
Contig ID
                  8013 1.R1040
                  zhf700960701.h1
                  BLASTX
```

5'-most EST Method

NCBI GI q2914706 BLAST score 719



E value 7.0e-76 Match length 238 % identity 56

NCBI Description (AC003974) putative homeobox protein [Arabidopsis thaliana]

Seq. No. 7110

Contig ID 8013 2.R1040

5'-most EST LIB3109-044-Q1-K1-A5

Method BLASTX
NCBI GI g3298555
BLAST score 176
E value 1.0e-12
Match length 48
% identity 62

NCBI Description (AC004681) putative homeobox protein, 3' partial

[Arabidopsis thaliana]

Seq. No. 7111

Contig ID 8016 1.R1040

5'-most EST uC-gmropic065a06b1

Method BLASTX
NCBI GI g1669601
BLAST score 581
E value 6.0e-60
Match length 190
% identity 63

NCBI Description (D88747) AR401 [Arabidopsis thaliana]

Seq. No. 7112

Contig ID 8020_1.R1040

5'-most EST LIB3039-047-Q1-E1-D6

Method BLASTX
NCBI GI g3341443
BLAST score 330
E value 8.0e-31
Match length 115
% identity 57

NCBI Description (AJ223074) acid phosphatase [Glycine max]

Seq. No. 7113

Contig ID 8031_1.R1040 5'-most EST vwf700678476.h1

Seq. No. 7114

Contig ID 8044_1.R1040

5'-most EST LIB3039-047-Q1-E1-A7

Method BLASTX
NCBI GI 9731852
BLAST score 274
E value 3.0e-24
Match length 90
% identity 58

NCBI Description HYPOTHETICAL 48.3 KD PROTEIN IN MOB1-SGA1 INTERGENIC REGION

>gi_626372_pir__S48469 probable membrane protein YIL103w -

yeast (Saccharomyces cerevisiae) >gi_558703_emb_CAA86277_

(Z38125) orf, len: 425, CAI: 0.19 [Saccharomyces

cerevisiae]

7115 Seq. No. Contig ID 8048 1.R1040 5'-most EST LIB3039-047-Q1-E1-B11 Seq. No. 7116 Contig ID 8053 1.R1040 5'-most EST bth700843962.h1 Method BLASTX NCBI GI g3834325 BLAST score 1145 E value 1.0e-125 Match length 316 % identity 68 (AC005679) Strong similarity to gb AF067141 gamma-glutamyl NCBI Description hydrolase from Arabidopsis thaliana. ESTs gb_R83955, gb_T45062, gb_T22220, gb_AA586207, gb_AI099851 and gb AI00672 come from this gene. [Arabidopsis thaliana] 7117 Seq. No. Contig ID 8054 1.R1040 5'-most EST LIB3049-041-Q1-E1-B4 7118 Seq. No. Contig ID 8069 1.R1040 5'-most EST kl1701211702.hl Method BLASTX NCBI GI g3212112 BLAST score 355 E value 2.0e-33 Match length 163 % identity 44 NCBI Description (Y17394) prefoldin subunit 3 [Homo sapiens] >gi 4507873 ref NP 003363.1 pVBP1 von Hippel-Lindau binding protein Seq. No. 7119 Contig ID 8069 2.R1040 5'-most EST LIB3039-046-01-E1-G11 Method BLASTX NCBI GI q3915477 BLAST score 170 E value 4.0e-12 Match length 92 37 % identity HYPOTHETICAL 20.9 KD PROTEIN T06G6.9 IN CHROMOSOME I NCBI Description >gi 3879580 emb CAB04707 (Z81587) Similarity to Human VHL binding protein-1 (TR:Q15765); cDNA EST EMBL:C13429 comes

binding protein-1 (TR:Q15765); cDNA EST EMBL:C13429 comes from this gene; cDNA EST EMBL:C10456 comes from this gene; cDNA EST yk486h8.3 comes from this gene [Caenorhabditis

elegans]

Seq. No. Contig ID

7120

Contig ID 8069_3.R1040 5'-most EST sat701008639.h1

Method BLASTX NCBI GI g3915477



BLAST score 145 E value 2.0e-09 Match length 71 % identity 41

NCBI Description HYPOTHETICAL 20.9 KD PROTEIN T06G6.9 IN CHROMOSOME I

>gi_3879580_emb_CAB04707_ (Z81587) Similarity to Human VHL
binding protein-1 (TR:Q15765); cDNA EST EMBL:C13429 comes
from this gene; cDNA EST EMBL:C10456 comes from this gene;
cDNA EST yk486h8.3 comes from this gene [Caenorhabditis]

elegans]

Seq. No. 7121

Contig ID 8080 1.R1040

5'-most EST LIB3039-046-Q1-E1-E6

Method BLASTX
NCBI GI g1169199
BLAST score 634
E value 5.0e-66
Match length 171
% identity 70

NCBI Description DNA-DAMAGE-REPAIR/TOLERATION PROTEIN DRT102

>gi_479739_pir__S35271 hypothetical protein - Arabidopsis

thaliana >gi 166928 (L11368) [Arabidopsis thaliana

unidentified mRNA sequence, complete cds.], gene product

[Arabidopsis thaliana]

Seq. No. 7122

Contig ID 8088_1.R1040 5'-most EST jex700904890.h1

Method BLASTN

NCBI GI g3982595

BLAST score 209

E value 1.0e-114

Match length 221

% identity 99

NCBI Description Glycine max cationic peroxidase 2 (Prx2) mRNA, complete cds

Seq. No. 7123

Contig ID 8091 1.R1040

5'-most EST LIB3039-046-Q1-E1-C8

Seq. No. 7124

Contig ID 8093 1.R1040

5'-most EST LIB3049-012-Q1-E1-B2

Method BLASTX
NCBI GI g122087
BLAST score 673
E value 1.0e-70
Match length 136
% identity 100

NCBI Description HISTONE H3 >gi_81849_pir_S04520 histone H3 (clone pH3c-1)

- alfalfa >gi_82609 pir A26014 histone H3 - wheat >gi_19607 emb CAA31964 (X13673) histone H3 (AA 1-136) [Medicago sativa] >gi_19609 emb CAA31965 (X13674) histone H3 (AA 1-136) [Medicago sativa] >gi_21797 emb CAA25451 (X00937) H3 histone [Triticum aestivum] >gi_488565 (U09459) histone H3.1 [Medicago sativa] >gi_2565419 (AF026803)

histone H3 [Onobrychis viciifolia]

7125 Seq. No. Contig ID 8102 1.R1040 5'-most EST pmv700890295.h1 Method BLASTX NCBI GI g3021409 BLAST score 668 E value 5.0e-70 Match length 244 % identity 25 NCBI Description (Y12781) transducin (beta) like 1 protein [Homo sapiens] Seq. No. 7126 Contig ID 8111 1.R1040 5'-most EST LIB3039-045-Q1-E1-H1 Seq. No. 7127 Contig ID 8117 1.R1040 5'-most EST LIB3039-045-Q1-E1-H8 Seq. No. 7128 Contig ID 8122 1.R1040 5'-most EST sat701014331.h1 Method BLASTX NCBI GI q1871185 BLAST score 231 E value 8.0e-19 Match length 115 % identity 37 NCBI Description (U90439) seven in absentia isolog [Arabidopsis thaliana] Seq. No. 7129 Contig ID 8126 1.R1040 5'-most EST LIB3039-045-Q1-E1-F6 Method BLASTX NCBI GI g125722 BLAST score 284 E value 4.0e-25 Match length 176 % identity 47 KUNITZ-TYPE TRYPSIN INHIBITOR KTI1 PRECURSOR NCBI Description >gi_81814_pir__JQ1091 trypsin inhibitor KTil (Kunitz) soybean >gi_256635_bbs_115028 (S45035) Kunitz trypsin inhibitor KTil [soybeans, Peptide, 203 aa] [Glycine max] Seq. No. 7130 Contig ID 8127 1.R1040 5'-most EST leu701149923.h1 Method BLASTX NCBI GI q4007792 BLAST score 246 E value 9.0e-21 Match length 65 % identity 71 NCBI Description (AL034463) Xenopus 14s cohesin smc1 subunit homolog

[Schizosaccharomyces pombe]

Seq. No. 7131 Contig ID 8129 1.R1040 5'-most EST q5676949 Method BLASTN NCBI GI q303838 BLAST score 72 E value 7.0e-32 Match length 208 % identity 84

NCBI Description Rice mRNA for 40S subunit ribosomal protein, complete cds

Seq. No. 7132

Contig ID 8129 2.R1040

5'-most EST jC-gmfl02220076f03d1

Seq. No. 7133

Contig ID 8129 3.R1040

5'-most EST jC-gmst02400007d04d1

Method BLASTX
NCBI GI g2129648
BLAST score 160
E value 2.0e-10
Match length 189
% identity 34

NCBI Description MYB-related protein 33,3K - Arabidopsis thaliana

>gi_1263095_emb_CAA90809_ (Z54136) MYB-related protein

[Arabidopsis thaliana]

Seq. No. 7134

Contig ID 8129 4.R1040

5'-most EST LIB3040-034-Q1-E1-F5

Method BLASTX
NCBI GI g548852
BLAST score 368
E value 3.0e-35
Match length 82
% identity 82

NCBI Description 40S RIBOSOMAL PROTEIN S21 >gi_481227_pir__S38357 ribosomal

protein S21 - rice >gi_303839_dbj_BAA02158_ (D12633) 40S

subunit ribosomal protein [Oryza sativa]

Seq. No. 7135

Contig ID 8131_1.R1040 5'-most EST g4405656

Seq. No. 7136

Contig ID 8136_1.R1040 5'-most EST ssr700557227.h1

Seq. No. 7137

Contig ID 8136_2.R1040 5'-most EST kl1701213934.h1

Seq. No. 7138

Contig ID 8136_3.R1040 5'-most EST xpa700792412.h1

```
Seq. No.
                   7139
Contig ID
                   8139 1.R1040
5'-most EST
                   LIB3051-030-Q1-K1-D10
Seq. No.
                   7140
Contig ID
                   8147 1.R1040
5'-most EST
                   rrt700645919.h1
Seq. No.
                   7141
Contig ID
                   8150 1.R1040
5'-most EST
                   trc700565861.h1
Seq. No.
                   7142
Contig ID
                   8151 1.R1040
5'-most EST
                   ssr700556967.hl
Method
                   BLASTX
NCBI GI
                   g3150402
                   1̃329
BLAST score
E value
                   1.0e-147
Match length
                   339
% identity
                   79
NCBI Description
                   (AC004165) putative malonyl-CoA: Acyl carrier protein
                   transacylase [Arabidopsis thaliana]
Seq. No.
                   7143
                   8151 2.R1040
Contig ID
5'-most EST
                   gsv7\overline{0}1052805.h1
Method
                   BLASTX
NCBI GI
                   g2347187
BLAST score
                   239
E value
                   6.0e-20
Match length
                   90
% identity
                   60
NCBI Description
                   (AC002338) putative malonyl-CoA:Acyl carrier protein
                   transacylase, 3' partial [Arabidopsis thaliana]
Seq. No.
                   7144
Contig ID
                   8151 3.R1040
5'-most EST
                   trc700567853.h1
Seq. No.
                   7145
Contig ID
                   8153 1.R1040
5'-most EST
                   LIB3087-010-Q1-K1-D4
Method
                   BLASTX
NCBI GI
                   g1350720
BLAST score
                   456
E value
                   2.0e-45
Match length
                   111
% identity
                  77
NCBI Description
                  60S RIBOSOMAL PROTEIN L32
Seq. No.
                  7146
Contig ID
                  8154 2.R1040
5'-most EST
```

Seq. No. 7147

LIB3039-045-Q1-E1-C5

5'-most EST

Method

NCBI GI

```
Contig ID
                   8158 1.R1040
5'-most EST
                   LIB3040-030-Q1-E1-E5
Method
                   BLASTX
                   g122003
NCBI GI
BLAST score
                   404
E value
                   4.0e-39
Match length
                   115
% identity
                   72
NCBI Description HISTONE H2A >gi_82089_pir__JQ1182 histone H2A.1 - tomato
Seq. No.
                   7148
Contig ID
                   8173 1.R1040
5'-most EST
                   LIB3039-021-Q1-E1-H8
Method
                   BLASTX
NCBI GI
                   g2262151
BLAST score
                   374
E value
                   2.0e-35
Match length
                   198
% identity
                   46
NCBI Description
                  (AC002330) predicted protein of unknown function
                   [Arabidopsis thaliana]
Seq. No.
                   7149
Contig ID
                   8173 2.R1040
5'-most EST
                   bth700846442.h1
Seq. No.
                   7150
Contig ID
                   8178 1.R1040
5'-most EST
                   jex700909126.h1
Seq. No.
                   7151
Contig ID
                   8189 1.R1040
5'-most EST
                   LIB3039-044-Q1-E1-F4
Method
                   BLASTX
NCBI GI
                   g4314388
BLAST score
                   475
E value
                   1.0e-47
Match length
                   129
% identity
                   69
                  (AC006232) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7152
                   8189 2.R1040
Contig ID
5'-most EST
                  sat701006251.h2
Method
                  BLASTX
NCBI GI
                  q4314388
BLAST score
                  394
E value
                  3.0e-38
Match length
                  117
% identity
                  64
NCBI Description
                  (AC006232) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  7153
                  8198 1.R1040
Contig ID
```

1324

LIB3039-044-Q1-E1-G6

BLASTX

g2281094



BLAST score 508 E value 2.0e-51 Match length 132 % identity 71

NCBI Description (AC002333) molybdenum cofactor biosynthesis protein E isolog [Arabidopsis thaliana] >gi_4469121 emb CAB38428

(AJ133519) molybdopterin synthase large subunit

[Arabidopsis thaliana]

Seq. No. 7154

Contig ID 8204_1.R1040 5'-most EST kmv700738093.h1

Method BLASTX
NCBI GI g2339980
BLAST score 380
E value 2.0e-36
Match length 164
% identity 55

NCBI Description (Y11337) RGA2 protein [Arabidopsis thaliana]

Seq. No. 7155

Contig ID 8204_2.R1040 5'-most EST jex700904794.h1

Method BLASTX
NCBI GI g2339978
BLAST score 247
E value 5.0e-21
Match length 100
% identity 60

NCBI Description (Y11336) RGA1 protein [Arabidopsis thaliana]

Seq. No. 7156

Contig ID 8217 1.R1040

5'-most EST LIB3039-044-Q1-E1-D10

Seq. No. 7157

Contig ID 8223_1.R1040 5'-most EST pmv700893311.h1

Method BLASTX
NCBI GI g3421123
BLAST score 1060
E value 1.0e-116
Match length 227
% identity 85

NCBI Description (AF043538) 20S proteasome beta subunit PBG1 [Arabidopsis

thaliana]

Seq. No. 7158

Contig ID 8226_1.R1040 5'-most EST epx701109794.h1

Method BLASTX
NCBI GI g2832686
BLAST score 240
E value 6.0e-20
Match length 138
% identity 46

NCBI Description (AL021712) putative protein [Arabidopsis thaliana]



```
Seq. No.
                   7159
Contig ID
                   8226 2.R1040
5'-most EST
                  uC-gmflminsoy053e06b1
Seq. No.
Contig ID
                   8236 1.R1040
5'-most EST
                  LIB3039-043-Q1-E1-H8
Method
                  BLASTX
NCBI GI
                  g3122861
BLAST score
                  157
E value
                   3.0e-10
Match length
                  114
% identity
                   34
NCBI Description
                  D-3-PHOSPHOGLYCERATE DEHYDROGENASE (PGDH) >gi 2649798
                   (AE001048) phosphoglycerate dehydrogenase (serA)
                   [Archaeoglobus fulgidus]
Seq. No.
                   7161
Contig ID
                   8240 1.R1040
5'-most EST
                  LIB3039-041-Q1-E1-E11
Method
                  BLASTX
NCBI GI
                  g299370
BLAST score
                  216
E value
                   4.0e-17
Match length
                  118
% identity
                   42
NCBI Description
                  (S58039) zeta-crystallin/quinone reductase [human, liver,
                   Peptide, 328 aa] [Homo sapiens]
                   >gi_4503067_ref_NP_001880.1_pCRYZ_ crystallin, zeta
                   (quinone reductase)
Seq. No.
                  7162
Contig ID
                  8245 1.R1040
5'-most EST
                  LIB3039-008-Q1-E1-A7
Method
                  BLASTX
                  g3108209
NCBI GI
BLAST score
                  809
                  1.0e-86
E value
Match length
                  172
% identity
                   85
NCBI Description
                  (AF028809) eukaryotic cap-binding protein [Arabidopsis
                  thaliana]
                  7163
Seq. No.
Contig ID
                  8245 2.R1040
5'-most EST
                  g5058009
Method
                  BLASTX
```

5'-most EST g5058009

Method BLASTX

NCBI GI g3108209

BLAST score 298

E value 6.0e-27

Match length 62

% identity NCBI Description

(AF028809) eukaryotic cap-binding protein [Arabidopsis

thaliana]

85

7164

Seq. No.



```
8246 1.R1040
Contig ID
5'-most EST
                   LIB3040-008-Q1-E1-A8
Seq. No.
                   7165
                   8249 1.R1040
Contig ID
5'-most EST
                   LIB3039-043-Q1-E1-G12
                   BLASTX
Method
                   g3309170
NCBI GI
BLAST score
                   269
                   3.0e-23
E value
                   130
Match length
                   45
% identity
                  (AF071314) COP9 complex subunit 4 [Mus musculus]
NCBI Description
                   7166
Seq. No.
Contig ID
                   8249 2.R1040
5'-most EST
                   LIB3109-005-Q1-K1-H3
                   BLASTX
Method
NCBI GI
                   g3309170
BLAST score
                   721
                   4.0e-76
E value
Match length
                   284
% identity
                   53
                  (AF071314) COP9 complex subunit 4 [Mus musculus]
NCBI Description
                   7167
Seq. No.
Contig ID
                   8258 1.R1040
5'-most EST
                   LIB3039-043-Q1-E1-D6
                   BLASTX
Method
NCBI GI
                   g4432835
BLAST score
                   406
                   3.0e-39
E value
Match length
                   141
                   52
% identity
NCBI Description
                  (AC006283) unknown protein [Arabidopsis thaliana]
                   7168
Seq. No.
                   8263_1.R1040
Contig ID
                   zsg7\overline{0}1127333.h1
5'-most EST
Method
                   BLASTX
                   g232031
NCBI GI
                   589
BLAST score
                   1.0e-60
E value
Match length
                   224
% identity
                   54
                   ELONGATION FACTOR 1 BETA' >gi_322851_pir__S29224
translation elongation factor eEF-1 beta' chain - rice
NCBI Description
                   >gi_218161_dbj_BAA02253_ (D12821) elongation factor 1 beta'
                   [Oryza sativa]
Seq. No.
                   7169
                   8263 3.R1040
Contig ID
```

5'-most EST

smc700749004.h1

Method BLASTX
NCBI GI g232031
BLAST score 273
E value 6.0e-24



Match length 76 % identity 66

NCBI Description ELONGATION FACTOR 1 BETA' >gi 322851 pir S29224

translation elongation factor eEF-1 beta chain - rice >gi_218161_dbj_BAA02253_ (D12821) elongation factor 1 beta

[Oryza sativa]

Seq. No. 7170

Contig ID 8263_4.R1040

5'-most EST jC-gmro02800025g06d1

Method BLASTN
NCBI GI g398607
BLAST score 35
E value 3.0e-10
Match length 59
% identity 90

NCBI Description A.thaliana mRNA for elongation factor 1 beta

Seq. No. 7171

Contig ID 8269_1.R1040 5'-most EST eep700867383.h1

Method BLASTX
NCBI GI g2559012
BLAST score 1412
E value 1.0e-157
Match length 373
% identity 73

NCBI Description (AF026293) chaperonin containing t-complex polypeptide 1,

beta subunit; CCT-beta [Homo sapiens] >gi 4090929

(AF026166) chaperonin-containing TCP-1 beta subunit homolog

[Homo sapiens]

Seq. No. 7172

Contig ID 8278_1.R1040 5'-most EST ncj700979022.h1

Method BLASTX
NCBI GI g3618214
BLAST score 215
E value 8.0e-17
Match length 189
% identity 31

NCBI Description (AL031579) dihydrofolate reductase [Schizosaccharomyces

pombe]

Seq. No. 7173

Contig ID 8287 1.R1040

5'-most EST uC-gmrominsoy113h05b1

Method BLASTX
NCBI GI g2129742
BLAST score 295
E value 1.0e-26
Match length 75
% identity 73

NCBI Description stress-induced protein OZI1 precursor - Arabidopsis

thaliana >gi_790583 (U20347) mRNA corresponding to this gene accumulates in response to ozone stress and pathogen (bacterial) infection; putative pathogenesis-related



protein [Arabidopsis thaliana] >gi_2252869 (AF013294) No
definition line found [Arabidopsis thaliana]

Seq. No. 7174

Contig ID 8287_2.R1040 5'-most EST gsv701050369.h1

Method BLASTX
NCBI GI g2129742
BLAST score 302
E value 2.0e-27
Match length 75
% identity 73

NCBI Description stress-induced protein OZI1 precursor - Arabidopsis

thaliana >gi_790583 (U20347) mRNA corresponding to this gene accumulates in response to ozone stress and pathogen (bacterial) infection; putative pathogenesis-related protein [Arabidopsis thaliana] >gi_2252869 (AF013294) No

definition line found [Arabidopsis thaliana]

Seq. No. 7175

Contig ID 8287 3.R1040

5'-most EST LIB3040-050-Q1-E1-F7

Method BLASTX
NCBI GI g2129742
BLAST score 295
E value 1.0e-26
Match length 75
% identity 72

NCBI Description stress-induced protein OZI1 precursor - Arabidopsis

thaliana >gi_790583 (U20347) mRNA corresponding to this gene accumulates in response to ozone stress and pathogen (bacterial) infection; putative pathogenesis-related protein [Arabidopsis thaliana] >gi 2252869 (AF013294) No

definition line found [Arabidopsis thaliana]

Seq. No. 7176

Contig ID 8287_4.R1040

5'-most EST uC-gmflminsoy035e03b1

Method BLASTX
NCBI GI g2129742
BLAST score 299
E value 3.0e-27
Match length 75
% identity 73

NCBI Description stress-induced protein OZI1 precursor - Arabidopsis

thaliana >gi_790583 (U20347) mRNA corresponding to this gene accumulates in response to ozone stress and pathogen (bacterial) infection; putative pathogenesis-related protein [Arabidopsis thaliana] >gi_2252869 (AF013294) No

definition line found [Arabidopsis thaliana]

Seq. No. 7177

Contig ID 8287_5.R1040 5'-most EST g4313658 Method BLASTX NCBI GI g2129742 BLAST score 283



E value 4.0e-25 Match length 75 % identity 69

NCBI Description

stress-induced protein OZI1 precursor - Arabidopsis thaliana >gi_790583 (U20347) mRNA corresponding to this gene accumulates in response to ozone stress and pathogen (bacterial) infection; putative pathogenesis-related protein [Arabidopsis thaliana] >gi_2252869 (AF013294) No definition line found [Arabidopsis thaliana]

Seq. No. 7178

Contig ID 8289_1.R1040 5'-most EST leu701154286.h1

Seq. No. 7179

Contig ID 8292_1.R1040 5'-most EST leu701151247.h1

Method BLASTX
NCBI GI g2529662
BLAST score 534
E value 2.0e-54
Match length 108
% identity 94

NCBI Description (AC002535) putative small nuclear ribonucleoprotein, Sm D2

[Arabidopsis thaliana] >gi_3738278 (AC005309) putative small nuclear ribonucleoprotein, Sm D2 [Arabidopsis

thaliana]

Seq. No. 7180

Contig ID 8292_2.R1040

5'-most EST LIB3072-008-Q1-E1-F11

Method BLASTX
NCBI GI g2529662
BLAST score 533
E value 2.0e-54
Match length 108
% identity 94

NCBI Description (AC002535) putative small nuclear ribonucleoprotein, Sm D2

[Arabidopsis thaliana] >gi_3738278 (AC005309) putative small nuclear ribonucleoprotein, Sm D2 [Arabidopsis

thaliana]

Seq. No. 7181

Contig ID 8292 3.R1040

5'-most EST LIB3040-057-Q1-E1-G6

Method BLASTX
NCBI GI g2529662
BLAST score 405
E value 1.0e-39
Match length 87
% identity 90

NCBI Description (AC002535) putative small nuclear ribonucleoprotein, Sm D2

[Arabidopsis thaliana] >gi_3738278 (AC005309) putative small nuclear ribonucleoprotein, Sm D2 [Arabidopsis

thaliana]

Seq. No. 7182



```
8294 1.R1040
Contig ID
                   LIB3039-043-Q1-E1-A4
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3759184
BLAST score
                   165
E value
                   1.0e-11
Match length
                   66
                   47
% identity
                  (AB018441) phi-1 [Nicotiana tabacum]
NCBI Description
Seq. No.
                   7183
Contig ID
                   8295 1.R1040
5'-most EST
                   leu701155106.h1
Method
                   BLASTX
NCBI GI
                   q232024
```

Method BLASTX
NCBI GI g232024
BLAST score 163
E value 2.0e-10
Match length 247
% identity 34
NCBI Description PROTEIN

NCBI Description PROTEIN E6 >gi_421806_pir__A46130 fiber protein - upland cotton >gi_2129498_pir__S65061 fiber protein E6 (clone CKE6-1A) - upland cotton >gi_167323 (M92051) 5' start site is putative; putative [Gossypium hirsutum] >gi_1000084

(U30505) E6 [Gossypium hirsutum]

Seq. No. 7184

Contig ID 8295 2.R1040

5'-most EST jC-gmro02910020g09a1

Seq. No. 7185

Contig ID 8300_1.R1040

5'-most EST LIB3039-043-Q1-E1-B2

Method BLASTX
NCBI GI 94097547
BLAST score 171
E value 8.0e-12
Match length 77
% identity 35

NCBI Description (U64906) ATFP3 [Arabidopsis thaliana]

Seq. No. 7186

Contig ID 8301 1.R1040

5'-most EST LIB3039-043-Q1-E1-B4

Method BLASTX
NCBI GI g3367593
BLAST score 312
E value 3.0e-28
Match length 129
% identity 45

NCBI Description (AL031135) putative protein [Arabidopsis thaliana] >gi_3805841_emb_CAA21461_ (AL031986) putative protein

[Arabidopsis thaliana]

Seq. No. 7187

Contig ID 8304 1.R1040

5'-most EST LIB3039-042-Q1-E1-G5



Seq. No. 7188

Contig ID 8309_1.R1040 5'-most EST g5677411

Seq. No. 7189

Contig ID 8318_1.R1040

5'-most EST LIB3049-043-Q1-E1-C1

Method BLASTX
NCBI GI g1707074
BLAST score 253
E value 2.0e-21
Match length 209
% identity 34

NCBI Description (U80450) M01E11.2 [Caenorhabditis elegans]

Seq. No. 7190

Contig ID 8325 1.R1040

5'-most EST LIB3039-021-Q1-E1-H6

Method BLASTX
NCBI GI g1169782
BLAST score 437
E value 4.0e-43
Match length 149
% identity 63

NCBI Description FUSCA PROTEIN FUS6 >gi_432446 (L26498) FUS6 [Arabidopsis

thaliana]

Seq. No. 7191

Contig ID 8331 1.R1040

5'-most EST jC-gmro02910029g06a1

Seq. No. 7192

Contig ID 8331_2.R1040 5'-most EST g5607075

Seq. No. 7193

Contig ID 8331_3.R1040 5'-most EST vzy700753608.h1

Seq. No. 7194

Contig ID 8331 4.R1040

5'-most EST LIB3039-042-Q1-E1-C7

Seq. No. 7195

Contig ID 8340_1.R1040 5'-most EST pmv700889160.h1

Method BLASTX
NCBI GI g2832642
BLAST score 407
E value 5.0e-39
Match length 463
% identity 29

NCBI Description (AL021710) putative protein [Arabidopsis thaliana]

Seq. No. 7196

Contig ID 8344 1.R1040

5'-most EST LIB3107-031-Q1-K1-C4



```
BLASTX
Method
NCBI GI
                   g2982283
BLAST score
                   382
E value
                   2.0e-36
Match length
                  148
% identity
                   51
                  (AF051226) PREG-like protein [Picea mariana]
NCBI Description
                   7197
Seq. No.
Contig ID
                   8344 2.R1040
5'-most EST
                  LIB3107-072-Q1-K1-C8
                   7198
Seq. No.
Contig ID
                   8349 1.R1040
5'-most EST
                   LIB3170-011-Q1-J1-E1
Method
                   BLASTX
NCBI GI
                   g459009
BLAST score
                   236
E value
                   1.0e-19
Match length
                   79
% identity
                   54
                  (U00037) similar to multifunctional aminoacyl-tRNA
NCBI Description
                   synthetase, especially to the prolyl-tRNA synthetase region
                   [Caenorhabditis elegans]
Seq. No.
                   7199
Contig ID
                   8349 2.R1040
5'-most EST
                   LIB3072-025-Q1-E1-A6
Method
                   BLASTX
                   g459009
NCBI GI
                   211
BLAST score
                   8.0e-17
E value
Match length
                   80
                   49
% identity
                   (U00037) similar to multifunctional aminoacyl-tRNA
NCBI Description
                   synthetase, especially to the prolyl-tRNA synthetase region
                   [Caenorhabditis elegans]
                   7200
Seq. No.
                   8349 3.R1040
Contig ID
5'-most EST
                   taw700657102.hl
                   7201
Seq. No.
                   8353 1.R1040
Contig ID
5'-most EST
                   asn7\overline{0}1131479.h1
Method
                   BLASTX
NCBI GI
                   g2911060
                   154
BLAST score
E value
                   5.0e-10
Match length
                   88
```

Seq. No. 7202

% identity

NCBI Description

Contig ID 8361 1.R1040

29

[Arabidopsis thaliana]

(AL021961) putative protein [Arabidopsis thaliana]

>gi 3297826 emb CAA19884.1 (AL031032) putative protein



5'-most EST LIB3039-041-Q1-E1-F12

Seq. No. 7203

Contig ID 8370_1.R1040 5'-most EST ncj700981540.h1

Method BLASTX
NCBI GI g3335337
BLAST score 446
E value 5.0e-44
Match length 119
% identity 76

NCBI Description (AC004512) Similar to acyl carrier protein, mitochondrial

precursor (ACP) NADH-ubiquinone oxidoreductase 9.6 KD subunit (MYACP-1), gb_L23574 from A. thaliana. ESTs gb_Z30712, gb_Z30713, gb_Z26204, gb_N37975 and gb_N96330

come from this gene

Seq. No. 7204

Contig ID 8370_2.R1040

5'-most EST LIB3039-033-Q1-E1-E11

Method BLASTX
NCBI GI g3335337
BLAST score 327
E value 2.0e-30
Match length 115
% identity 65

NCBI Description (AC004512) Similar to acyl carrier protein, mitochondrial

precursor (ACP) NADH-ubiquinone oxidoreductase 9.6 KD subunit (MYACP-1), gb_L23574 from A. thaliana. ESTs gb_Z30712, gb_Z30713, gb_Z26204, gb_N37975 and gb_N96330

£.

come from this gene

Seq. No. 7205

Contig ID 8376 1.R1040

5'-most EST LIB3049-018-Q1-E1-D1

Method BLASTX
NCBI GI g3135265
BLAST score 354
E value 3.0e-33
Match length 101
% identity 68

NCBI Description (AC003058) unknown protein [Arabidopsis thaliana]

Seq. No. 7206

Contig ID 8376_2.R1040

5'-most EST LIB3109-034-Q1-K1-B11

Method BLASTX
NCBI GI g3135265
BLAST score 441
E value 1.0e-43
Match length 101
% identity 81

NCBI Description (AC003058) unknown protein [Arabidopsis thaliana]

Seq. No. 7207

Contig ID 8378 1.R1040

5'-most EST LIB3\overline{1}09-006-Q1-K1-B12



Method BLASTX NCBI GI q746510 BLAST score 454 1.0e-44 E value Match length 213 % identity 43 (U23517) similar to ubiquitin conjugating enzyme NCBI Description [Caenorhabditis elegans] 7208 Seq. No. Contig ID 8378 3.R1040 5'-most EST LIB3072-010-Q1-E1-D9 Method BLASTX NCBI GI q746510 BLAST score 325 E value 3.0e-30 Match length 97 % identity 57 (U23517) similar to ubiquitin conjugating enzyme NCBI Description [Caenorhabditis elegans] Seq. No. 7209 Contig ID 8378 5.R1040 5'-most EST LIB3138-007-Q1-N1-A10

Seq. No. 7210

Contig ID 8378_6.R1040 5'-most EST zzp700831760.h1

Method BLASTX
NCBI GI 9746510
BLAST score 171
E value 2.0e-12
Match length 44
% identity 64

NCBI Description (U23517) similar to ubiquitin conjugating enzyme

[Caenorhabditis elegans]

Seq. No. 7211

Contig ID 8387_1.R1040

5'-most EST LIB3039-041-Q1-E1-D6

Method BLASTX
NCBI GI g3461813
BLAST score 382
E value 2.0e-40
Match length 125
% identity 66

NCBI Description (AC004138) putative sucrose/H+ symporter [Arabidopsis

thaliana]

Seq. No. 7212

Contig ID 8397 1.R1040

5'-most EST LIB3\overline{106}-020-Q1-K1-B1

Seq. No. 7213

Contig ID 8408_1.R1040 5'-most EST ncj700984619.h1

Method BLASTX

% identity

55



```
g2586127
NCBI GI
                  309
BLAST score
E value
                  4.0e-28
Match length
                  124
% identity
                  52
                  (U89510) b-keto acyl reductase [Hordeum vulgare]
NCBI Description
                  7214
Seq. No.
                  8408 2.R1040
Contig ID
                  uxk7\overline{0}0668489.h1
5'-most EST
Seq. No.
                  7215
                  8409 1.R1040
Contig ID
                  smw700646216.h1
5'-most EST
Seq. No.
                  7216
                   8413 1.R1040
Contig ID
                  jC-qmle01810041a04a1
5'-most EST
                  7217
Seq. No.
Contig ID
                   8413 2.R1040
                  LIB3039-018-Q1-E1-H4
5'-most EST
                   7218
Seq. No.
                   8417 1.R1040
Contig ID
                   LIB3074-029-Q1-K1-A6
5'-most EST
                   BLASTX
Method
                   g1173456
NCBI GI
                   328
BLAST score
                   3.0e-30
E value
                   97
Match length
                   58
% identity
NCBI Description SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3 (SNRNP CORE PROTEIN
                   D3) (SM-D3) >gi_600750 (U15009) Sm D3 [Homo sapiens]
                   7219
Seq. No.
                   8417_2.R1040
Contig ID
                   LIB3170-062-Q1-J1-A1
5'-most EST
                   BLASTX
Method
                   g2708715
NCBI GI
BLAST score
                   142
                   1.0e-08
E value
Match length
                   46
                   54
% identity
                  (AF038598) small nuclear ribonucleoprotein Sm D3
NCBI Description
                   [Drosophila melanogaster]
                   7220
Seq. No.
Contig ID
                   8422 1.R1040
                   trc700562809.hl
5'-most EST
Method
                   BLASTX
                   q1877026
NCBI GI
BLAST score
                   202
E value
                   1.0e-20
Match length
                   86
```

NCBI Description (D78336) ribosomal protein S19 [Oryza sativa]



7221 Seq. No.

8427 1.R1040 Contig ID

LIB3094-078-Q1-K1-A11 5'-most EST

Method BLASTX NCBI GI q4508069 230 BLAST score 7.0e-19 E value 161 Match length 35 % identity

(AC005882) 12246 [Arabidopsis thaliana] NCBI Description

7222 Seq. No.

8429 1.R1040 Contig ID $leu7\overline{0}1145719.h1$ 5'-most EST

Method BLASTX NCBI GI q730456 BLAST score 615 7.0e-64 E value 140 Match length % identity 80

NCBI Description 40S RIBOSOMAL PROTEIN S19

7223 Seq. No.

8438 1.R1040 Contig ID wvk700681268.hl 5'-most EST

Method BLASTX NCBI GI q2493694 BLAST score 299 E value 6.0e-27 119 Match length % identity 53

PHOTOSYSTEM II REACTION CENTRE W PROTEIN PRECURSOR (PSII NCBI Description

6.1 KD PROTEIN) >gi 1076268_pir__S53025 photosystem II protein - spinach >gi_728716 emb_CAA59409 (X85038) protein

of photosystem II [Spinacia oleracea]

7224 Seq. No.

8438 2.R1040 Contig ID leu701147974.h1 5'-most EST

Seq. No. 7225

8438 3.R1040 Contig ID

LIB3073-017-Q1-K1-G7 5'-most EST

BLASTX Method g2493694 NCBI GI 214 BLAST score 4.0e-17 E value 116 Match length % identity 40

NCBI Description PHOTOSYSTEM II REACTION CENTRE W PROTEIN PRECURSOR (PSII

6.1 KD PROTEIN) >gi_1076268_pir__S53025 photosystem II protein - spinach >gi_728716_emb_CAA59409_ (X85038) protein

of photosystem II [Spinacia oleracea]

Seq. No. 7226

8441 1.R1040 Contig ID

5'-most EST

Method



```
vwf700675988.h1
5'-most EST
                  7227
Seq. No.
Contig ID
                  8446 1.R1040
5'-most EST
                  fC-gmro700866877d3
Method
                  BLASTX
NCBI GI
                  g1169533
BLAST score
                  155
                  3.0e-10
E value
                  37
Match length
% identity
                  86
                  ENOLASE (2-PHOSPHOGLYCERATE DEHYDRATASE)
NCBI Description
                  (2-PHOSPHO-D-GLYCERATE HYDRO-LYASE)
                  >gi_515827_emb_CAA56645_ (X80474) enolase [Neocallimastix
                  frontalis]
Seq. No.
                  7228
Contig ID
                  8464 1.R1040
                  LIB3049-034-Q1-E1-E11
5'-most EST
                  BLASTN
Method
                  q3927915
NCBI GI
BLAST score
                  135
                  1.0e-69
E value
                   311
Match length
% identity
NCBI Description Fagus sylvatica mRNA for glycine-rich protein
                   7229
Seq. No.
Contig ID
                   8464 2.R1040
                   uaw700663386.hl
5'-most EST
Method
                   BLASTX
NCBI GI
                   q1076626
                   498
BLAST score
                   2.0e-50
E value
                   92
Match length
                   98
% identity
                   glycine rich protein - common tobacco
NCBI Description
                   >gi 790473 emb CAA58702 (X83731) soluble, glycine rich
                   protein [Nicotiana tabacum]
                   7230
Seq. No.
                   8464 3.R1040
Contig ID
                   zpv700760105.hl
5'-most EST
                   BLASTX
Method
                   g1076626
NCBI GI
                   166
BLAST score
                   5.0e-16
E value
                   46
Match length
                   96
% identity
                   glycine rich protein - common tobacco
NCBI Description
                   >gi_790473_emb_CAA58702_ (X83731) soluble, glycine rich
                   protein [Nicotiana tabacum]
                   7231
Seq. No.
                   8464 4.R1040
Contig ID
```

awf700836556.hl

BLASTN

1338



NCBI GI g3927915 BLAST score 61 E value 7.0e-26 Match length 133 % identity 86

NCBI Description Fagus sylvatica mRNA for glycine-rich protein

Seq. No. 7232

Contig ID 8465 1.R1040

5'-most EST LIB3039-039-Q1-E1-G9

Method BLASTN
NCBI GI g3819163
BLAST score 283
E value 1.0e-158
Match length 335
% identity 96

NCBI Description Glycine max cctd gene

Seq. No. 7233

Contig ID 8465_2.R1040 5'-most EST zhf700956825.h1

Method BLASTN
NCBI GI g3819163
BLAST score 146
E value 2.0e-76
Match length 319
% identity 93

NCBI Description Glycine max cctd gene

Seq. No. 7234

Contig ID 8468_1.R1040

5'-most EST uC-gmropic023h03b1

Method BLASTX
NCBI GI g3935167
BLAST score 533
E value 3.0e-54
Match length 123
% identity 84

NCBI Description (AC004557) F17L21.10 [Arabidopsis thaliana]

Seq. No. 7235

Contig ID 8468_2.R1040 5'-most EST wrg700787076.h2

Method BLASTX
NCBI GI g3935167
BLAST score 520
E value 7.0e-53
Match length 123
% identity 83

NCBI Description (AC004557) F17L21.10 [Arabidopsis thaliana]

Seq. No. 7236

Contig ID 8473_1.R1040 5'-most EST bth700845213.h1

Method BLASTX NCBI GI g2979550 BLAST score 437



```
E value
                   6.0e-43
Match length
                   173
% identity
                   50
NCBI Description
                   (AC003680) putative 7-ethoxycoumarin O-deethylase
                   [Arabidopsis thaliana]
Seq. No.
                   7237
Contig ID
                   8475 1.R1040
5'-most EST
                   LIB3039-040-Q1-E1-A11
Method
                   BLASTX
NCBI GI
                   q3687250
BLAST score
                   361
E value
                   6.0e-34
Match length
                   241
% identity
                   38
NCBI Description
                   (AC005169) putative arginine n-methyltransferase
                   [Arabidopsis thaliana]
Seq. No.
                   7238
Contig ID
                   8476 1.R1040
5'-most EST
                   jC-gmro02910027d07a1
Method
                   BLASTX
NCBI GI
                   q4056503
BLAST score
                   180
E value
                   6.0e-13
Match length
                   36
% identity
                   92
NCBI Description
                  (AC005896) unknown protein [Arabidopsis thaliana]
Seq. No.
                   7239
                   8476 2.R1040
Contig ID
5'-most EST
                   LIB3049-052-Q1-E1-G5
Method
                   BLASTX
NCBI GI
                   q4056488
BLAST score
                   324
E value
                   2.0e-46
Match length
                   199
% identity
                   48
NCBI Description
                  (AC005896) unknown protein [Arabidopsis thaliana]
Seq. No.
                   7240
                   8476 3.R1040
Contig ID
5'-most EST
                   LIB3106-031-Q1-K1-A11
Seq. No.
                   7241
Contig ID
                   8476 4.R1040
5'-most EST
                   eep700865621.h1
Method
                  BLASTX
NCBI GI
                   g730826
```

Method BLASTX
NCBI GI g730826
BLAST score 218
E value 1.0e-17
Match length 97
% identity 31

NCBI Description SPLICING FACTOR, ARGININE/SERINE-RICH 4 (PRE-MRNA SPLICING FACTOR SRP75) >gi_1082713 pir A48133 pre-mRNA splicing

SRp75 - human >gi_307438 (L14076) pre-mRNA splicing factor

[Homo sapiens]

E value

5.0e-34

```
7242
Seq. No.
                  8476 5.R1040
Contig ID
5'-most EST
                  uaw700666680.hl
Method
                  BLASTX
                  g4056503
NCBI GI
                  180
BLAST score
                  3.0e-13
E value
Match length
                  36
% identity
                  92
                  (AC005896) unknown protein [Arabidopsis thaliana]
NCBI Description
                  7243
Seq. No.
                  8476 8.R1040
Contig ID
5'-most EST
                  gsv701054710.hl
                  7244
Seq. No.
Contig ID
                  8476 11.R1040
5'-most EST
                  hyd700727629.h1
Method
                  BLASTX
NCBI GI
                  q4056488
                  187
BLAST score
                  1.0e-28
E value
Match length
                  96
% identity
                  65
NCBI Description
                  (AC005896) unknown protein [Arabidopsis thaliana]
Seq. No.
                  7245
                  8483 1.R1040
Contig ID
5'-most EST
                  LIB3051-051-Q1-K1-F9
                  7246
Seq. No.
                  8483 2.R1040
Contig ID
5'-most EST
                  awf700842960.h1
                  7247
Seq. No.
                  8483 3.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy204g02b1
Seq. No.
                  7248
Contig ID
                  8487 1.R1040
                  LIB3039-039-Q1-E1-G5
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2809246
BLAST score
                  246
                   5.0e-21
E value
Match length
                  78
% identity
                   62
NCBI Description
                  (AC002560) F2401.15 [Arabidopsis thaliana]
                  7249
Seq. No.
                  8490 1.R1040
Contig ID
5'-most EST
                  LIB3039-039-Q1-E1-F2
Method
                  BLASTX
NCBI GI
                  g1065515
BLAST score
                   359
```

154 Match length % identity 47 (U40420) weak similarity to procollagen alpha chain 1(V) NCBI Description chain [Caenorhabditis elegans] 7250 Seq. No. 8492 1.R1040 Contig ID leu701144444.hl 5'-most EST BLASTX Method NCBI GI q2351374 BLAST score 533 2.0e-54 E value Match length 111 92 % identity (U54560) putative 26S proteasome subunit athMOV34 NCBI Description [Arabidopsis thaliana] 7251 Seq. No. 8493 1.R1040 Contig ID leu701149723.hl 5'-most EST Method BLASTX q3810594 NCBI GI BLAST score 456 2.0e-45 E value 123 Match length % identity 74 (AC005398) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 7252 8494 1.R1040 Contig ID LIB3039-039-Q1-E1-E4 5'-most EST BLASTX Method g2702274 NCBI GI 362 BLAST score

2.0e-34 E value 103 Match length % identity 67

(AC003033) unknown protein [Arabidopsis thaliana] NCBI Description

7253 Seq. No.

8497 1.R1040 Contig ID

LIB3039-039-Q1-E1-F1 5'-most EST

Method BLASTX g2894596 NCBI GI BLAST score 639 1.0e-66 E value 187 Match length % identity 67

(AL021889) putative protein [Arabidopsis thaliana] NCBI Description

7254 Seq. No.

8497 2.R1040 Contig ID trc700562114.h1 5'-most EST

Seq. No. 7255

8500 1.R1040 Contig ID

LIB3051-018-Q1-E1-H7 5'-most EST



```
7256
Seq. No.
                  8500 2.R1040
Contig ID
                  hrw701063012.hl
5'-most EST
                  7257
Seq. No.
                  8506 1.R1040
Contig ID
                  pcp700993831.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4220474
BLAST score
                  240
E value
                  4.0e-20
                  84
Match length
                  65
% identity
                  (AC006069) putative myosin heavy chain [Arabidopsis
NCBI Description
                  thaliana]
                  7258
Seq. No.
                  8511 1.R1040
Contig ID
                  LIB3039-039-Q1-E1-B11
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3913794
BLAST score
                  665
                  1.0e-69
E value
                  166
Match length
                  73
% identity
                  GLUTATHIONE PEROXIDASE 1 >gi 2326453 emb_CAA74775 (Y14429)
NCBI Description
                  glutathione peroxidase [Helianthus annuus]
Seq. No.
                  7259
Contig ID
                  8511 2.R1040
5'-most EST
                  uC-qmflminsoy045c07b1
Method
                  BLASTX
NCBI GI
                  q2760606
BLAST score
                  162
                  6.0e-28
E value
Match length
                  93
% identity
                  80
                   (AB001568) phospholipid hydroperoxide glutathione
NCBI Description
                  peroxidase-like protein [Arabidopsis thaliana] >gi_3004869
                   (AF030132) glutathione peroxidase; ATGP1 [Arabidopsis
                   thaliana] >gi_4539451_emb_CAB39931.1_ (AL049500)
                  phospholipid hydroperoxide glutathione peroxidase
                   [Arabidopsis thaliana]
                  7260
Seq. No.
                  8517 1.R1040
Contig ID
                  seb700651955.hl
5'-most EST
                  BLASTX
Method
                  g2431771
NCBI GI
```

Method BLASTX
NCBI GI g2431771
BLAST score 228
E value 2.0e-18
Match length 113
% identity 44

NCBI Description (U62753) acidic ribosomal protein P2b [Zea mays]

Seq. No. 7261



```
8518 1.R1040
Contig ID
                  uC-gmrominsoy187d07b1
5'-most EST
                  7262
Seq. No.
                  8520 1.R1040
Contig ID
                  leu701154316.h1
5'-most EST
                  BLASTX
Method
                  g1172977
NCBI GI
BLAST score
                  766
                  2.0e-81
E value
                  171
Match length
                  86
% identity
                  60S RIBOSOMAL PROTEIN L18 >gi_606970 (U15741) cytoplasmic
NCBI Description
                  ribosomal protein L18 [Arabidopsis thaliana]
                  7263
Seq. No.
                  8520 2.R1040
Contig ID
                  LIB3170-018-Q1-J1-A1
5'-most EST
Method
                  BLASTX
                  g1172977
NCBI GI
BLAST score
                  556
                  4.0e-57
E value
                  122
Match length
                  84
% identity
                  60S RIBOSOMAL PROTEIN L18 >gi 606970 (U15741) cytoplasmic
NCBI Description
                  ribosomal protein L18 [Arabidopsis thaliana]
                  7264
Seq. No.
                  8521 1.R1040
Contig ID
                  pxt700943611.h1
5'-most EST
Method
                  BLASTX
                  q2737973
NCBI GI
                   590
BLAST score
E value
                   4.0e-61
                   156
Match length
% identity
                   72
NCBI Description
                  (U83625) protein kinase ZmMEK1 [Zea mays]
                   7265
Seq. No.
                   8524 1.R1040
Contig ID
5'-most EST
                   LIB3040-058-Q1-E1-E4
                   BLASTX
Method
NCBI GI
                   g549732
BLAST score
                   211
E value
                   1.0e-16
Match length
                   79
% identity
                   HYPOTHETICAL 16.2 KD PROTEIN IN PIR3-APE2 INTERGENIC REGION
NCBI Description
                   >gi 481110 pir S37791 hypothetical protein YKL160w - yeast
                   (Saccharomyces cerevisiae) >gi_407488_emb_CAA81494_
                   (Z26877) unknown [Saccharomyces cerevisiae]
                   >gi_486279_emb_CAA82002_ (Z28160) ORF YKL160w
                   [Saccharomyces cerevisiae] >gi_1582545_prf__2118404F ORF
                   [Saccharomyces cerevisiae]
```

7266

8546 1.R1040

Seq. No.

Contig ID



5'-most EST LIB3074-024-Q1-E1-C12

Seq. No. 7267

Contig ID 8552 1.R1040

5'-most EST LIB3039-038-Q1-E1-D7

Seq. No. 7268

Contig ID 8558_1.R1040

5'-most EST LIB3039-038-Q1-E1-C7

Seq. No. 7269

Contig ID 8562_1.R1040

5'-most EST LIB3094-046-Q1-K1-H11

Method BLASTN
NCBI GI g499066
BLAST score 488
E value 0.0e+00
Match length 678
% identity 99

NCBI Description G.max gmrl gene

Seq. No. 7270

Contig ID 8565 1.R1040

5'-most EST jC-gmro02910006c04a1

Seq. No. 7271

Contig ID 8565_2.R1040 5'-most EST gsv701048378.h1

Seq. No. 7272

Contig ID 8570_1.R1040 5'-most EST taw700655350.h1

Seq. No. 7273

Contig ID 8579 1.R1040

5'-most EST jC-gmst02400075g11d1

Method BLASTX
NCBI GI g2367418
BLAST score 489
E value 3.0e-49
Match length 140
% identity 62

NCBI Description (AF000392) peptide transporter [Lotus japonicus]

Seq. No. 7274

Contig ID 8581 1.R1040

5'-most EST LIB3039-037-Q1-E1-G7

Seq. No. 7275

Contig ID 8584 1.R1040 5'-most EST gsv701050179.h1

Seq. No. 7276

Contig ID 8591_1.R1040 5'-most EST V4L-01-Q1-B1-C3

Method BLASTX NCBI GI g3451075



BLAST score 1683 E value 0.0e+00 Match length 432 % identity 70 NCBI Description (AL03132

NCBI Description (AL031326) putative protein [Arabidopsis thaliana]

Seq. No. 7277

Contig ID 8594_1.R1040

5'-most EST LIB3039-012-Q1-E1-D9

Method BLASTX
NCBI GI g2851506
BLAST score 323
E value 5.0e-30
Match length 82
% identity 70

NCBI Description DYSKERIN (NUCLEOLAR PROTEIN NAP57)

>gi_2739325_emb_CAA84402_ (Z34922) NAP57 [Rattus

norvegicus]

Seq. No. 7278

Contig ID 8594_2.R1040

5'-most EST LIB3051-034-Q1-K1-G7

Method BLASTX
NCBI GI g2851506
BLAST score 340
E value 5.0e-34
Match length 98
% identity 71

NCBI Description DYSKERIN (NUCLEOLAR PROTEIN NAP57)

>qi 2739325 emb CAA84402 (Z34922) NAP57 [Rattus

norvegicus]

Seq. No. 7279

Contig ID 8617_1.R1040 5'-most EST kl1701208274.h1

Seq. No. 7280

Contig ID 8618_1.R1040

5'-most EST LIB3049-024-Q1-E1-D10

Method BLASTX
NCBI GI g3176874
BLAST score 2036
E value 0.0e+00
Match length 733
% identity 71

NCBI Description (AF065639) cucumisin-like serine protease [Arabidopsis

thaliana]

Seq. No. 7281

Contig ID 8618_5.R1040 5'-most EST ncj700978984.h1

Method BLASTX
NCBI GI g3176874
BLAST score 225
E value 1.0e-18
Match length 62
% identity 61

1346



NCBI Description (AF065639) cucumisin-like serine protease [Arabidopsis thaliana]

Seq. No. 7282

Contig ID 8620_1.R1040

5'-most EST fC-gmse700752057a2

Method BLASTX
NCBI GI g1174470
BLAST score 820
E value 9.0e-88
Match length 264
% identity 58

NCBI Description OLIGOSACCHARYL TRANSFERASE STT3 SUBUNIT HOMOLOG (B5)

(INTEGRAL MEMBRANE PROTEIN 1) >gi_508543 (L34260) integral membrane protein 1 [Mus musculus] >gi_1588285_prf__2208301A

integral membrane protein [Mus musculus]

Seq. No. 7283

Contig ID 8620_2.R1040 5'-most EST smc700745470.h1

Seq. No. 7284

Contig ID 8630_1.R1040

5'-most EST LIB3039-036-Q1-E1-F10

Method BLASTX
NCBI GI g2464915
BLAST score 473
E value 1.0e-47
Match length 111
% identity 77

NCBI Description (Z99708) transcription initiation factor like protein

[Arabidopsis thaliana]

Seq. No. 7285

Contig ID 8632_1.R1040

5'-most EST LIB3040-010-Q1-E1-B9

Seq. No. 7286

Contig ID 8632_3.R1040

5'-most EST LIB3106-074-Q1-K1-B10

Seq. No. 7287

Contig ID 8634_1.R1040 5'-most EST leu701150274.h1

Method BLASTX
NCBI GI 9445613
BLAST score 967
E value 1.0e-105
Match length 237
% identity 78

NCBI Description ribosomal protein L7 [Solanum tuberosum]

Seq. No. 7288

Contig ID 8639 1.R1040 5'-most EST crh700851304.h1

Seq. No. 7289

```
8645 1.R1040
Contig ID
                  LIB3040-043-Q1-E1-E12
5'-most EST
                  BLASTN
Method
                  q166739
NCBI GI
BLAST score
                   126
E value
                   2.0e-64
                   314
Match length
                   85
% identity
NCBI Description A.thaliana histone H4 gene, complete cds
Seq. No.
                   7290
                   8647 1.R1040
Contig ID
                   zhf700954010.h1
5'-most EST
                   7291
Seq. No.
Contig ID
                   8647 2.R1040
5'-most EST
                   LIB3039-036-Q1-E1-B2
                   7292
Seq. No.
Contig ID
                   8649 1.R1040
                   LIB3093-047-Q1-K1-H7
5'-most EST
                   BLASTX
Method
                   g3434986
NCBI GI
                   430
BLAST score
                   2.0e-42
E value
                   118
Match length
% identity
                   31
                   (AB016895) Pop3 [Schizosaccharomyces pombe]
NCBI Description
Seq. No.
                   7293
                   8650 1.R1040
Contig ID
                   uC-gmronoir031b08b1
5'-most EST
                   BLASTX
Method
                   g4510347
NCBI GI
                   232
BLAST score
                   3.0e-19
E value
                   72
Match length
                   71
% identity
                   (AC006921) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   7294
Seq. No.
                   8654 1.R1040
Contig ID
                   uC-gmronoir058f08b1
5'-most EST
                   BLASTX
Method
                   g3953458
NCBI GI
                   482
BLAST score
                   2.0e-48
E value
                   150
Match length
                   69
 % identity
                   (AC002328) F20N2.3 [Arabidopsis thaliana]
NCBI Description
                   7295
 Seq. No.
                   8654 2.R1040
 Contig ID
                   LIB3092-041-Q1-K1-H11
 5'-most EST
```

BLASTX Method

NCBI GI q3953458 360 BLAST score



E value 3.0e-34
Match length 116
% identity 67

NCBI Description (AC002328) F20N2.3 [Arabidopsis thaliana]

Seq. No. 7296

Contig ID 8655_1.R1040 5'-most EST g5678154

Seq. No. 7297

Contig ID 8681 1.R1040

5'-most EST uC-gmrominsoy250g03b1

Method BLASTX
NCBI GI g2244799
BLAST score 398
E value 1.0e-38
Match length 136
% identity 57

NCBI Description (Z97336) carnitine racemase homolog [Arabidopsis thaliana]

Seq. No. 7298

Contig ID 8682 1.R1040 5'-most EST leu701157256.h1

Method BLASTX
NCBI GI g3152606
BLAST score 601
E value 5.0e-62
Match length 170
% identity 64

NCBI Description (AC004482) putative ring zinc finger protein [Arabidopsis

thaliana]

Seq. No. 7299

Contig ID 8682_4.R1040 5'-most EST kl1701207818.h1

Seq. No. 7300

Contig ID 8684 1.R1040

5'-most EST LIB3039-035-Q1-E1-F12

Seq. No. 7301

Contig ID 8685 1.R1040

5'-most EST jC-gmf102220103e03a1

Method BLASTX
NCBI GI g4567283
BLAST score 1561
E value 0.0e+00
Match length 523
% identity 69

NCBI Description (AC006841) unknown protein [Arabidopsis thaliana]

Seq. No. 7302

Contig ID 8685_2.R1040

5'-most EST uC-gmropic026b05b1

Method BLASTX
NCBI GI g4567283
BLAST score 779



```
7.0e-83
E value
Match length
                  285
% identity
                  58
NCBI Description
                  (AC006841) unknown protein [Arabidopsis thaliana]
Seq. No.
                  7303
Contig ID
                  8685 3.R1040
5'-most EST
                  LIB3051-013-Q1-E1-F6
                  7304
Seq. No.
Contig ID
                  8691 1.R1040
5'-most EST
                  LIB3072-014-Q1-E1-D8
                  7305
Seq. No.
Contig ID
                  8699 1.R1040
5'-most EST
                  LIB3039-035-Q1-E1-E2
Method
                  BLASTX
NCBI GI
                  g2827086
BLAST score
                  322
E value
                  6.0e-30
Match length
                  100
% identity
                  58
                  (AF022778) DNA recombination and repair protein [Homo
NCBI Description
                   sapiens] >gi_3328152 (AF073362) endo/exonuclease Mre11
                   [Homo sapiens]
                  7306
Seq. No.
                   8701 1.R1040
Contig ID
5'-most EST
                  LIB3039-005-Q1-E1-B3
Seq. No.
                  7307
Contig ID
                  8702 1.R1040
                  LIB3039-035-Q1-E1-E5
5'-most EST
                  BLASTX
Method
                   g4335759
NCBI GI
BLAST score
                  165
                   2.0e-11
E value
                   98
Match length
% identity
                   47
                  (AC006284) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7308
Seq. No.
                   8706 1.R1040
Contig ID
5'-most EST
                   leu701155875.h1
Seq. No.
                   7309
                   8707_1.R1040
Contig ID
5'-most EST
                  LIB3039-035-Q1-E1-C12
Method
                   BLASTX
                   g2832629
NCBI GI
```

BLAST score 153 7.0e-10 E value Match length 116 % identity

(AL021711) 4-coumarate-CoA ligase - like [Arabidopsis NCBI Description

thaliana]



```
7310
Seq. No.
                  8709 1.R1040
Contig ID
5'-most EST
                  qsv701047702.hl
Seq. No.
                  7311
Contig ID
                  8714 1.R1040
5'-most EST
                  kl1701211410.hl
                  BLASTX
Method
                  q1076427
NCBI GI
BLAST score
                  570
E value
                  6.0e-59
                  115
Match length
                  89
% identity
                  ubiquitin--protein ligase (EC 6.3.2.19) - Arabidopsis
NCBI Description
                  thaliana
                  7312
Seq. No.
Contig ID
                  8716 1.R1040
                  LIB3049-055-Q1-E1-H4
5'-most EST
Method
                  BLASTN
                  q3860314
NCBI GI
BLAST score
                  207
                   1.0e-112
E value
                   345
Match length
% identity
                   92
                  Cicer arietinum mRNA for 40S ribosomal protein S19, partial
NCBI Description
Seq. No.
                   8716 2.R1040
Contig ID
5'-most EST
                   LIB3051-032-Q1-K1-G5
Method
                   BLASTN
NCBI GI
                   g3860314
BLAST score
                   123
                   1.0e-62
E value
Match length
                   275
% identity
                   86
NCBI Description Cicer arietinum mRNA for 40S ribosomal protein S19, partial
                   7314
Seq. No.
                   8717 1.R1040
Contig ID
                   LIB3106-019-Q1-K1-C1
5'-most EST
Seq. No.
                   7315
                   8720 1.R1040
Contig ID
                   LIB3039-035-Q1-E1-A9
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g2281092
BLAST score 233
E value 3.0e-19
Match length 68
% identity 72

NCBI Description (AC002333) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7316

Contig ID 8721_1.R1040 5'-most EST zhf700958240.h1



```
Seq. No.
                    7317
```

8721 2.R1040 Contig ID 5'-most EST sat701009738.h1

Seq. No.

7318

8723 1.R1040 Contig ID

5'-most EST

LIB3039-035-Q1-E1-A11

Method BLASTN NCBI GI g4220422 BLAST score 106 E value 2.0e-52 Match length 226 87 % identity

NCBI Description Fragaria x ananassa putative cellulase mRNA, complete cds

Seq. No.

7319

Contig ID

8727 1.R1040

5'-most EST

LIB3170-043-Q1-K1-D6

BLASTX Method q730030 NCBI GI BLAST score 157 E value 1.0e-09 Match length 279 % identity 12

NCBI Description

MICRONUCLEAR LINKER HISTONE POLYPROTEIN (MIC LH) [CONTAINS:

LINKER HISTONE PROTEINS ALPHA, BETA, DELTA AND GAMMA]

>gi 161804 (M87306) micronuclear linker histone polyprotein

[Tetrahymena thermophila]

7320 Seq. No.

8727 2.R1040 Contig ID 5'-most EST g4284398 Method BLASTX NCBI GI g1066104 BLAST score 147 E value 6.0e-09 Match length 236

12 % identity

NCBI Description (U39735) high molecular weight basic nuclear protein

[Pleuronectes americanus]

Seq. No. 7321

8740 1.R1040 Contig ID 5'-most EST epx701105109.h1

BLASTX Method NCBI GI q4455172 BLAST score 836 E value 2.0e-89 Match length 228 % identity 70

NCBI Description (AL035521) putative protein [Arabidopsis thaliana]

Seq. No.

7322

Contig ID 8742 1.R1040

5'-most EST uC-gmrominsoy086a04b1

Method BLASTX NCBI GI q4337210



BLAST score 371 3.0e-35 E value 135 Match length 53 % identity

NCBI Description (AC006403) hypothetical protein [Arabidopsis thaliana]

7323 Seq. No.

8742 2.R1040 Contig ID

jC-gmf102220097b01d1 5'-most EST

7324 Seq. No.

8742 3.R1040 Contig ID wvk700680217.h2 5'-most EST

7325 Seq. No.

8743 1.R1040 Contig ID $hyd7\overline{0}0726342.h1$ 5'-most EST

7326 Seq. No.

8743 2.R1040 Contig ID

LIB3138-029-Q1-N1-F10 5'-most EST

7327 Seq. No.

8744 1.R1040 Contig ID

LIB3106-106-Q1-K1-F9 5'-most EST

7328 Seq. No.

8748 1.R1040 Contig ID ncj700985843.hl 5'-most EST

BLASTX Method q4567251 NCBI GI 682 BLAST score 2.0e-71 E value 303 Match length % identity

(AC007070) unknown protein [Arabidopsis thaliana] NCBI Description

7329

Seq. No. 8748 2.R1040 Contig ID g5342502 5'-most EST Method BLASTX g4567251 NCBI GI BLAST score 538 1.0e-54 E value Match length 154

68 % identity

(AC007070) unknown protein [Arabidopsis thaliana] NCBI Description

7330 Seq. No.

8748 5.R1040 Contig ID gsv701051285.hl 5'-most EST

7331 Seq. No.

8748 6.R1040 Contig ID ncj700985348.hl 5'-most EST

BLASTX Method NCBI GI g4567251



BLAST score 243 E value 9.0e-21 77 Match length 58 % identity

(AC007070) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

7332

Contig ID 5'-most EST

8754 1.R1040 LIB3039-034-Q1-E1-E7

Seq. No.

7333

Contig ID 5'-most EST 8759 1.R1040 leu701153853.h1

Seq. No.

7334

Contig ID 5'-most EST 8764 1.R1040

Method

LIB3039-016-Q1-E1-C8 BLASTN

NCBI GI

g3281847

BLAST score E value

42 2.0e-14

Match length % identity

86 87

NCBI Description

Arabidopsis thaliana DNA chromosome 4, BAC clone F28M20

(ESSAII project)

Seq. No.

7335

Contig ID 5'-most EST 8764 2.R1040 fua701040759.h1

Method

BLASTN g3281847

NCBI GI BLAST score

42

E value

2.0e-14

Match length

86

% identity

87

NCBI Description

Arabidopsis thaliana DNA chromosome 4, BAC clone F28M20

(ESSAII project)

Seq. No.

7336

Contig ID 5'-most EST 8765 1.R1040

Method

leu701149120.h1 BLASTX

NCBI GI

g2739279

BLAST score

535

E value

2.0e-54

Match length

208 53

% identity NCBI Description

(AJ223177) short chain alcohol dehydrogenase [Nicotiana

tabacum] >gi_2791348_emb_CAA11154_ (AJ223178) short chain

alcohol dehydrogenase [Nicotiana tabacum]

Seq. No.

7337

Contig ID 5'-most EST 8767 1.R1040 zsq701124165.h1

Seq. No.

7338

Contig ID

8777 1.R1040

```
5'-most EST
                    LIB3039-034-Q1-E1-B2
 Method
                    BLASTX
 NCBI GI
                    a3845099
 BLAST score
                    169
 E value
                    1.0e-11
Match length
                    124
 % identity
NCBI Description
                   (AE001373) predicted secreted protein [Plasmodium
                   falciparum]
Seq. No.
                   7339
Contig ID
                   8777 2.R1040
5'-most EST
                   LIB3049-022-Q1-E1-D4
Seq. No.
                   7340
Contig ID
                   8777 4.R1040
5'-most EST
                   LIB3073-004-Q1-K1-A6
Seq. No.
                   7341
Contig ID
                   8780 1.R1040
5'-most EST
                   LIB3051-082-Q1-K1-D2
Method
                   BLASTN
NCBI GI
                   q2696018
BLAST score
                   108
E value
                   1.0e-53
Match length
                   332
% identity
                   83
NCBI Description
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
                   MXC9, complete sequence [Arabidopsis thaliana]
Seq. No.
                   7342
Contig ID
                   8783 1.R1040
5'-most EST
                   epx701106758.h1
Seq. No.
                   7343
Contig ID
                   8784 1.R1040
5'-most EST
                   zhf700961867.h1
Method
                   BLASTX
NCBI GI
                   q3023956
BLAST score
                   191
E value
                   3.0e-14
Match length
                   168
% identity
NCBI Description
                  VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1 >gi_607003
                   (L28125) beta transducin-like protein [Podospora anserina]
Seq. No.
                   7344
Contig ID
                   8792 1.R1040
5'-most EST
                  LIB3039-033-Q1-E1-G5
Method
                  BLASTX
NCBI GI
                  g1181589
BLAST score
                  284
E value
                  2.0e-25
Match length
```

1355

(D83070) high mobility group protein [Canavalia gladiata] >gi_1483173_dbj_BAA13133_ (D86594) high mobility group

120

54

% identity

NCBI Description



protein [Canavalia gladiata]

```
7345
Seq. No.
                  8802_1.R1040
Contig ID
5'-most EST
                  g4301037
                  7346
Seq. No.
                  8802 3.R1040
Contig ID
                  LIB3039-013-Q1-E1-F5
5'-most EST
Seq. No.
                   7347
                   8806 1.R1040
Contig ID
                   LIB3049-008-Q1-E1-E4
5'-most EST
                   BLASTX
Method
                   g2407800
NCBI GI
BLAST score
                   509
E value
                   2.0e-51
Match length
                   110
% identity
                  (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
NCBI Description
                   7348
Seq. No.
                   8806 2.R1040
Contig ID
                   LIB3049-056-Q1-E1-B10
5'-most EST
                   BLASTX
Method
                   g2407800
NCBI GI
BLAST score
                   373
                   7.0e-36
E value
                   83
Match length
% identity
                   92
                   (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
NCBI Description
                   7349
Seq. No.
                   8815 1.R1040
Contig ID
                   epx701107574.hl
5'-most EST
                   BLASTN
Method
                   g432488
NCBI GI
                   138
BLAST score
                   3.0e-71
E value
                   282
Match length
                   87
 % identity
                   Wheat initiation factor 1A (eIF-1A) mRNA
NCBI Description
                   7350
 Seq. No.
                   8815 2.R1040
 Contig ID
                   fua701041538.h1
 5'-most EST
                   BLASTN
Method
                   g432488
NCBI GI
                   170
BLAST score
                   2.0e-90
 E value
Match length
                   282
 % identity
NCBI Description Wheat initiation factor 1A (eIF-1A) mRNA
```

Seq. No.

Contig ID 8815_3.R1040

5'-most EST LIB3049-017-Q1-E1-H7

7351

1356



```
BLASTN
Method
                   g432488
NCBI GI
BLAST score
                   135
E value
                   2.0e-69
                   282
Match length
                   87
% identity
NCBI Description Wheat initiation factor 1A (eIF-1A) mRNA
                   7352
Seq. No.
Contig ID
                   8815 4.R1040
                   LIB3051-053-Q1-K2-H1
5'-most EST
                   BLASTN
Method
NCBI GI
                   g432488
BLAST score
                   137
                   5.0e-71
E value
Match length
                   217
% identity
                    91
                   Wheat initiation factor 1A (eIF-1A) mRNA
NCBI Description
                   7353
Seq. No.
                   8815 6.R1040
Contig ID
5'-most EST
                    zsq701118418.h1
                   BLASTN
Method
                    g432488
NCBI GI
BLAST score
                    50
E value
                    3.0e-19
Match length
                    74
% identity
                    92
NCBI Description Wheat initiation factor 1A (eIF-1A) mRNA
                    7354
Seq. No.
                    8821 1.R1040
Contig ID
                    LIB3039-033-Q1-E1-B7
5'-most EST
                    BLASTX
Method
                    g4539005
NCBI GI
                    295
BLAST score
                    1.0e-26
E value
                    132
Match length
% identity
                    46
                    (ALO49481) putative oxidoreductase [Arabidopsis thaliana]
NCBI Description
                    7355
Seq. No.
                    8825 1.R1040
Contig ID
                    fua701040082.h1
5'-most EST
Method
                    BLASTX
                    q1345933
NCBI GI
                    1398
BLAST score
E value
                    1.0e-155
Match length
                    296
                    89
% identity
                    CITRATE SYNTHASE, GLYOXYSOMAL PRECURSOR (GCS) >gi_1084323_pir__S53007 citrate synthase - cucurbit
NCBI Description
                    >qi 975633 dbj BAA07328 (D38132) glyoxysomal citrate
                    synthase [Cucurbita sp.]
```

Seq. No. 7356

Contig ID 8828_1.R1040

NCBI Description

domestica]



```
LIB3039-039-Q1-E1-D2
5'-most EST
                  BLASTX
Method
                  g2500378
NCBI GI
                  409
BLAST score
                  7.0e-40
E value
                  83
Match length
                  89
% identity
NCBI Description 60S RIBOSOMAL PROTEIN L37
                  7357
Seq. No.
                  8828 2.R1040
Contig ID
                  LIB3074-039-Q1-K1-F8
5'-most EST
                  BLASTX
Method
                  g2500378
NCBI GI
                   437
BLAST score
                   5.0e-43
E value
                   95
Match length
                   83
% identity
NCBI Description 60S RIBOSOMAL PROTEIN L37
                   7358
Seq. No.
                   8835 1.R1040
Contig ID
                   fua701042366.h1
5'-most EST
                   BLASTX
Method
                   g1304227
NCBI GI
                   633
BLAST score
                   6.0e-66
E value
                   149
Match length
% identity
                   (D63781) Epoxide hydrolase [Glycine max]
NCBI Description
                   >qi 2764804 emb CAA55293 (X78547) epoxide hydrolase
                   [Glycine max]
                   7359
Seq. No.
                   8840 1.R1040
Contig ID
                   LIB3109-047-Q1-K1-A12
 5'-most EST
                   BLASTX
Method
                   g4519539
NCBI GI
 BLAST score
                   1526
                   1.0e-170
 E value
 Match length
                   364
                   79
 % identity
                   (AB016256) NAD-dependent sorbitol dehydrogenase [Malus
 NCBI Description
                   domestica]
                   7360
 Seq. No.
                   8840 2.R1040
 Contig ID
                   asn701134864.h2
 5'-most EST
                   BLASTX
 Method
                    g4519539
 NCBI GI
                    230
 BLAST score
                    1.0e-36
 E value
                    104
 Match length
                    76
 % identity
```

(AB016256) NAD-dependent sorbitol dehydrogenase [Malus

E value

Match length

6.0e-27

88



```
7361
Seq. No.
                  8842 1.R1040
Contig ID
                  LIB3049-012-Q1-E1-H3
5'-most EST
                  BLASTX
Method
NCBI GI
                  q4432837
BLAST score
                  268
                  3.0e-23
E value
                  148
Match length
                   47
% identity
NCBI Description (AC006283) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  8856 1.R1040
Contig ID
                  hyd700726103.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3928150
BLAST score
                   491
E value
                   2.0e-49
                   113
Match length
% identity
NCBI Description (AJ131049) hypothetical protein [Cicer arietinum]
                   7363
Seq. No.
                   8866 1.R1040
Contig ID
                   LIB3039-032-Q1-E1-D9
5'-most EST
                   BLASTX
Method
NCBI GI
                   q478318
BLAST score
                   174
                   3.0e-12
E value
Match length
                   100
% identity
                   36
                   immunophilin p59 - mouse >gi_410499_emb_CAA50231_ (X70887)
NCBI Description
                   p59 immunophilin [Mus musculus]
                   7364
Seq. No.
                   8871 1.R1040
Contig ID
                   LIB3039-032-Q1-E1-E7
5'-most EST
                   7365
Seq. No.
                   8873_1.R1040
Contig ID
                   leu701148713.h1
5'-most EST
                   BLASTX
Method
                   g2842486
NCBI GI
                   390
BLAST score
                   1.0e-37
E value
Match length
                   120
% identity
                   (AL021749) putative protein [Arabidopsis thaliana]
NCBI Description
                   7366
Seq. No.
                   8873 2.R1040
Contig ID
                   LIB3039-032-Q1-E1-F1
5'-most EST
                   BLASTX
Method
                   q4204306
NCBI GI
BLAST score
                   296
```

% identity NCBI Description (AC003027) 1cl prt seq No definition line found [Arabidopsis thaliana] 7367 Seq. No. 8875 1.R1040 Contig ID 5'-most EST LIB3039-032-Q1-E1-C2 Method BLASTX NCBI GI g2921209 BLAST score 499 2.0e-50 E value Match length 143 73 % identity NCBI Description (AF026148) beta-ketoacyl-ACP synthase I [Perilla frutescens] Seq. No. 7368 Contig ID 8883 1.R1040 5'-most EST LIB3109-001-Q1-K1-G11 Method BLASTX q2827536 338

٠٠٠.

NCBI GI BLAST score E value 5.0e-31 Match length 176 % identity 41 NCBI Description

(AL021633) hypothetical protein [Arabidopsis thaliana]

7369 Seq. No. Contig ID

8883 3.R1040

5'-most EST LIB3107-006-Q1-K1-A7

7370 Seq. No.

Contig ID 8883 4.R1040

5'-most EST LIB3052-016-Q1-N1-E3

Seq. No. 7371

8891 1.R1040 Contig ID 5'-most EST leu701157334.h1

Seq. No. 7372

8892 1.R1040 Contig ID 5'-most EST gsv701052888.h1

BLASTX Method NCBI GI g3881780 BLAST score 176 3.0e-12E value Match length 116 % identity 37

(Z48638) similar to lipid transfer protein [Caenorhabditis NCBI Description

elegans]

Seq. No. 7373

Contig ID 8901 1.R1040

LIB3039-004-Q1-E1-G1 5'-most EST

Method BLASTX NCBI GI g1183961 BLAST score 477

```
2.0e-57
 E value
                    205
 Match length
                    57
 % identity
 NCBI Description (Z69382) RNA elicase [Saccharomyces cerevisiae]
                    7374
 Seq. No.
                    8914 1.R1040
 Contig ID
                    LIB3053-001-Q1-B1-H12
 5'-most EST
                    7375
 Seq. No.
                    8922 1.R1040
 Contig ID
                    LIB3106-041-Q1-K1-E3
 5'-most EST
                    BLASTX
 Method
                    g3426039
 NCBI GI
                    315
 BLAST score
                    5.0e-29
 E value
                    91
Match length
 % identity
 NCBI Description (AC005168) unknown protein [Arabidopsis thaliana]
                    7376
 Seq. No.
                    8923 1.R1040
 Contig ID
                    LIB3138-104-Q1-N1-G7
 5'-most EST
                    BLASTX
 Method
                    q266936
 NCBI GI
                    527
 BLAST score
                    2.0e-53
 E value
                    143
 Match length
                    72
 % identity
                    50S RIBOSOMAL PROTEIN L27, CHLOROPLAST PRECURSOR (CL27)
 NCBI Description
                    >gi_282960_pir__A42840 ribosomal protein L27 - common
                    tobacco >gi_170306 (M98473) ribosomal protein L27
                     [Nicotiana Tabacum] >gi_170326 (M75731) ribosomal protein
                    L27 [Nicotiana tabacum]
                     7377
  Seq. No.
  Contig ID
                     8924 1.R1040
                    LIB3\overline{0}50-013-Q1-E1-G4
  5'-most EST
  Method
                     BLASTX
                     q2979565
  NCBI GI
  BLAST score
                     475
                     2.0e-47
  E value
  Match length
                     120
                     77
  % identity
                     (AC003680) putative sin3 associated polypeptide (SAP18)
  NCBI Description
                     [Arabidopsis thaliana]
                     7378
  Seq. No.
                     8931_1.R1040
  Contig ID
                     jex700905931.hl
  5'-most EST
                     BLASTX
  Method
                     q3650030
  NCBI GI
                     291
  BLAST score
                     5.0e-26
  E value
                     114
  Match length
                     58
  % identity
  NCBI Description (AC005396) unknown protein [Arabidopsis thaliana]
```

A STATE OF THE STA

```
7379
Seq. No.
                    8931 2.R1040
Contig ID
5'-most EST
                    LIB3039-031-Q1-E1-D3
Method
                    BLASTX
                    q3650030
NCBI GI
                    317
BLAST score
                    3.0e-29
E value
                    133
Match length
% identity
                    55
NCBI Description (AC005396) unknown protein [Arabidopsis thaliana]
                    7380
Seq. No.
                     8941 1.R1040
Contig ID
5'-most EST
                    LIB3170-005-Q1-K1-E2
                    7381
Seq. No.
                     8941 2.R1040
Contig ID
                    LIB3039-031-Q1-E1-B12
5'-most EST
                     BLASTX
Method
                     g3859570
NCBI GI
                     150
BLAST score
                     3.0e-13
E value
                     87
Match length
% identity
                     53
NCBI Description (AF098753) unknown [Oryza sativa]
                     7382
Seq. No.
                     8953 1.R1040
Contig ID
                     ncj7\overline{0}0982909.h1
5'-most EST
                     BLASTX
Method
                     g1170543
NCBI GI
                     679
BLAST score
                     3.0e-71
E value
                     244
Match length
                     57
% identity
                     DIHYDROXY-ACID DEHYDRATASE PRECURSOR (DAD) (2,3-DIHYDROXY
NCBI Description
                     ACID HYDROLYASE) >gi_1077091_pir_ S55205 dihydroxy-acid dehydratase (EC 4.2.1.9) - yeast (Saccharomyces cerevisiae) >gi_854590_emb_CAA60939_ (X87611) dihydroxyacid dehydratase
                     [Saccharomyces cerevisiae] >gi_1015650_emb_CAA89540_
                     (Z49516) ORF YJR016c [Saccharomyces cerevisiae]
                     7383
Seq. No.
                     8957 1.R1040
Contig ID
                     LIB3072-055-Q1-K1-E12
 5'-most EST
                     BLASTX
Method
                     q2827655
NCBI GI
BLAST score
                     185
                     1.0e-13
E value
Match length
                     128
```

41 % identity

NCBI Description (AL021637) hypothetical protein [Arabidopsis thaliana]

7384 Seq. No.

8957 2.R1040 Contig ID

LIB3094-080-Q1-K1-A9 5'-most EST

% identity

NCBI Description

45



```
BLASTX
Method
NCBI GI
                  g2827655
                  175
BLAST score
E value
                  5.0e-12
Match length
                  208
                  33
% identity
                  (AL021637) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  7385
Seq. No.
                  8958 1.R1040
Contig ID
5'-most EST
                  epx701109691.hl
                  BLASTX
Method
                  g3204129
NCBI GI
BLAST score
                  502
E value
                  1.0e-50
Match length
                  133
% identity
                  77
                  (AJ006768) histone H2A [Cicer arietinum]
NCBI Description
Seq. No.
                  7386
Contig ID
                  8958 2.R1040
                  LIB3074-040-Q1-K1-B2
5'-most EST
Method
                  BLASTX
                  q3776566
NCBI GI
BLAST score
                  398
E value
                  1.0e-38
                  106
Match length
% identity
                  75
                  (AC005388) Strong similarity to histone H2A gb AJ006768
NCBI Description
                   from Cicer arietinum. [Arabidopsis thaliana]
                   7387
Seq. No.
                   8978 1.R1040
Contig ID
                  LIB3039-030-Q1-E1-E7
5'-most EST
                  BLASTX
Method
                   g140474
NCBI GI
BLAST score
                   251
E value
                   7.0e-21
Match length
                   203
% identity
                   31
                  HYPOTHETICAL 69.2 KD PROTEIN IN HSP30-PMP1 INTERGENIC
NCBI Description
                   REGION >gi_83212_pir__S19434 probable transport protein
                   YCR023c - yeast (Saccharomyces cerevisiae)
                   >gi 1907167 emb CAA42315 (X59720) YCR023c, len:611
                   [Saccharomyces cerevisiae]
                   7388
Seq. No.
                   8978 2.R1040
Contig ID
5'-most EST
                   leu701153306.h1
Method
                   BLASTX
NCBI GI
                   q140474
BLAST score
                   159
E value
                   1.0e-10
Match length
                   67
```

1363

HYPOTHETICAL 69.2 KD PROTEIN IN HSP30-PMP1 INTERGENIC

REGION >gi 83212_pir__S19434 probable transport protein



YCR023c - yeast (Saccharomyces cerevisiae) >gi_1907167_emb_CAA42315_ (X59720) YCR023c, len:611 [Saccharomyces cerevisiae]

Seq. No. 7389

Contig ID 8978 3.R1040

5'-most EST uC-gmrominsoy172a11b1

Seq. No. 7390

Contig ID 8978_4.R1040 5'-most EST g5753013

Seq. No. 7391

Contig ID 8978_5.R1040 5'-most EST fua701041256.h1

Method BLASTX
NCBI GI g140474
BLAST score 178
E value 7.0e-13
Match length 76
% identity 43

NCBI Description HYPOTHETICAL 69.2 KD PROTEIN IN HSP30-PMP1 INTERGENIC

REGION >gi_83212_pir__S19434 probable transport protein

YCR023c - yeast (Saccharomyces cerevisiae)

>gi_1907167_emb_CAA42315 (X59720) YCR023c, len:611

[Saccharomyces cerevisiae]

Seq. No. 7392

Contig ID 8979 1.R1040

5'-most EST LIB3109-037-Q1-K1-A3

Method BLASTX
NCBI GI g1903364
BLAST score 392
E value 9.0e-38
Match length 154
% identity 58

NCBI Description (AC000104) EST gb T45093 comes from this gene. [Arabidopsis

thaliana

Seq. No. 7393

Contig ID 8983 1.R1040

5'-most EST LIB3039-030-Q1-E1-F2

Method BLASTX
NCBI GI g3395436
BLAST score 387
E value 3.0e-37
Match length 133
% identity 54

NCBI Description (AC004683) unknown protein [Arabidopsis thaliana]

Seq. No.

7394

Contig ID 8999_2.R1040

5'-most EST jC-gmst02400056c01a1

Method BLASTX
NCBI GI g3560264
BLAST score 170
E value 6.0e-12

E SERVICE SE

Match length 110 % identity 39

NCBI Description (AL031535) soll family protein [Schizosaccharomyces pombe]

Seq. No. 7395

Contig ID 8999 3.R1040

5'-most EST jC-qmf102220144f04a1

Seq. No. 7396

Contig ID 9002_1.R1040 5'-most EST leu701147244.h1

Seq. No. 7397

Contig ID 9004 1.R1040

5'-most EST uC-gmrominsoy276f10b1

Method BLASTN
NCBI GI g517257
BLAST score 74
E value 2.0e-33
Match length 130
% identity 89

NCBI Description Z.mays MNBla mRNA for DNA-binding protein

Seq. No. 7398

Contig ID 9010_1.R1040 5'-most EST vzy700751182.h1

Seq. No. 7399

Contig ID 9013_1.R1040

5'-most EST LIB3039-030-Q1-E1-A12

Seq. No. 7400

Contig ID 9024 1.R1040

5'-most EST LIB3039-037-Q1-E1-F12

Method BLASTN
NCBI GI g1806145
BLAST score 196
E value 1.0e-106
Match length 532
% identity 84

NCBI Description M.sativa mRNA for cdc2 kinase homologue, cdc2MsF

Seq. No. 7401

Contig ID 9024 2.R1040

5'-most EST LIB3170-034-Q1-K1-F1

Method BLASTX
NCBI GI g1806146
BLAST score 624
E value 3.0e-65
Match length 135
% identity 93

NCBI Description (X97317) cdc2MsF [Medicago sativa]

Seq. No. 7402

Contig ID 9025 1.R1040

5'-most EST LIB3051-030-Q1-K1-E12



Seq. No. 7403

Contig ID 9025_2.R1040 5'-most EST uaw700664486.h1

Seq. No. 7404

Contig ID 9034_1.R1040 5'-most EST zzp700832030.h1

Method BLASTX
NCBI GI g4406814
BLAST score 495
E value 1.0e-49
Match length 113
% identity 82

NCBI Description (AC006201) putative peptidyl-prolyl isomerase [Arabidopsis

thaliana]

Seq. No. 7405

Contig ID 9034_2.R1040 5'-most EST seb700653725.h1

Method BLASTN
NCBI GI g531828
BLAST score 41
E value 1.0e-13
Match length 81
% identity 88

NCBI Description Cloning vector pSport1, complete cds

Seq. No. 7406

Contig ID 9040 1.R1040

5'-most EST LIB3039-029-Q1-E1-D10

Seq. No. 7407

Contig ID 9042_1.R1040

5'-most EST fC-gmf1700905086i1

Seq. No. 7408

Contig ID 9042_2.R1040

5'-most EST LIB3039-029-Q1-E1-D12

Seq. No. 7409

Contig ID 9043 1.R1040

5'-most EST LIB3039-029-Q1-E1-D2

Method BLASTX
NCBI GI g4539470
BLAST score 358
E value 1.0e-33
Match length 123
% identity 66

NCBI Description (AL049500) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7410

Contig ID 9049_1.R1040 5'-most EST gsv701051392.h1

Method BLASTX
NCBI GI g4544399
BLAST score 274
E value 6.0e-24

1366



Match length 102 % identity 54

NCBI Description (AC007047) putative beta-ketoacyl-CoA synthase [Arabidopsis

thaliana]

Seq. No. 7411

Contig ID 9050_1.R1040 5'-most EST yuv700862829.h1

Method BLASTN
NCBI GI g536891
BLAST score 107
E value 9.0e-53
Match length 264
% identity 88

NCBI Description Wheat mRNA for protein H2A, complete cds, clone wcH2A-4

Seq. No. 7412

Contig ID 9050 2.R1040 5'-most EST leu701144726.h1

Method BLASTN
NCBI GI g536891
BLAST score 94
E value 4.0e-45
Match length 274
% identity 86

NCBI Description Wheat mRNA for protein H2A, complete cds, clone wcH2A-4

Seq. No. 7413

Contig ID 9050 3.R1040

5'-most EST LIB3040-021-Q1-E1-C2

Method BLASTN
NCBI GI g536891
BLAST score 99
E value 4.0e-48
Match length 285
% identity 86

NCBI Description Wheat mRNA for protein H2A, complete cds, clone wcH2A-4

Seq. No. 7414

Contig ID 9050 4.R1040

5'-most EST jC-gmle01810087g06a1

Method BLASTX
NCBI GI g3269284
BLAST score 281
E value 7.0e-25
Match length 56
% identity 100

NCBI Description (AL030978) histone H2A- like protein [Arabidopsis thaliana]

Seq. No. 7415

Contig ID 9053 1.R1040

5'-most EST LIB3039-029-Q1-E1-B5

Seq. No. 7416

Contig ID 9065_1.R1040

5'-most EST uC-gmropic040f08b1

Method BLASTN



g556421 NCBI GI BLAST score 56 2.0e-22 E value 120 Match length 96 % identity Stylosanthes humilis cinnamyl alcohol dehydrogenase (CAD1) NCBI Description mRNA, complete cds Seq. No. 7417 9065 2.R1040 Contig ID 5'-most EST gsv701051084.h1Method BLASTN g556421 NCBI GI BLAST score 52 E value 3.0e-20 Match length 109 94 % identity Stylosanthes humilis cinnamyl alcohol dehydrogenase (CAD1) NCBI Description mRNA, complete cds 7418 Seq. No. 9071 1.R1040 Contig ID 5'-most EST pmv700894565.h1 Seq. No. 7419 9076 1.R1040 Contig ID 5'-most EST LIB3106-011-Q1-K1-B5 BLASTX Method g3738329 NCBI GI BLAST score 322 1.0e-29 E value 88 Match length 78 % identity (AC005170) unknown protein [Arabidopsis thaliana] NCBI Description 7420 Seq. No. 9076 2.R1040 Contig ID ssr700557608.hl 5'-most EST 7421 Seq. No. 9079 1.R1040 Contig ID LIB3039-028-Q1-E1-H10 5'-most EST Method BLASTX

g4006872 NCBI GI BLAST score 231 4.0e-19 E value 60 Match length 73 % identity

NCBI Description (Z99707) methionyl aminopeptidase-like protein [Arabidopsis

thaliana

7422 Seq. No.

9080 1.R1040 Contig ID 5'-most EST bth700845178.h1

Method BLASTX g2708750 NCBI GI BLAST score 361



```
2.0e-34
E value
                  134
Match length
% identity
                  (AC003952) putative physical impedence protein [Arabidopsis
NCBI Description
                  thaliana]
                  7423
Seq. No.
                  9084 1.R1040
Contiq ID
                  LIB3039-039-Q1-E1-G8
5'-most EST
Seq. No.
                  7424
                  9084 2.R1040
Contig ID
                  LIB3170-088-Q1-K1-G1
5'-most EST
Seq. No.
                  7425
Contig ID
                  9089 1.R1040
                  LIB3073-019-Q1-K1-F10
5'-most EST
Method
                  BLASTX
                  g629858
NCBI GI
BLAST score
                  546
E value
                  9.0e-56
Match length
                  128
% identity
                  82
NCBI Description protein kinase C inhibitor - maize
Seq. No.
                  7426
                  9096 1.R1040
Contig ID
5'-most EST
                  LIB3039-028-Q1-E1-F8
Method
                  BLASTX
                   g3925703
NCBI GI
                  1007
BLAST score
                  1.0e-109
E value
                  242
Match length
% identity
                  81
                  (X95905) 14-3-3 protein [Lycopersicon esculentum]
NCBI Description
                   7427
Seq. No.
Contig ID
                   9096 2.R1040
5'-most EST
                   LIB3049-047-Q1-E1-C7
                  BLASTX
Method
                   g3023186
NCBI GI
                   360
BLAST score
E value
                   2.0e-34
Match length
                   91
                   77
% identity
                  14-3-3-LIKE PROTEIN 9 >gi 1771180 emb CAA67373.1 (X98865)
NCBI Description
                   14-3-3 protein [Lycopersicon esculentum]
                   7428
Seq. No.
                   9103 1.R1040
Contig ID
```

5'-most EST LIB3039-028-Q1-E1-D4

Seq. No.

7429

Contig ID 9104_1.R1040 5'-most EST sat701013255.h1

Method BLASTN NCBI GI g1150931



```
BLAST score
                  2.0e-15
E value
Match length
                  56
% identity
                  95
NCBI Description M.sativa mRNA for cyclin protein
                  7430
Seq. No.
                  9104 2.R1040
Contig ID
                  LIB3040-057-Q1-E1-C8
5'-most EST
Method
                  BLASTN
                  g1150931
NCBI GI
BLAST score
                  36
                  1.0e-10
E value
Match length
                  52
% identity
                  92
NCBI Description M.sativa mRNA for cyclin protein
Seq. No.
                  7431
                   9114 1.R1040
Contig ID
                   zzp700833874.h1
5'-most EST
Method
                  BLASTX
                   g3033397
NCBI GI
BLAST score
                  1699
E value
                  0.0e + 00
                   403
Match length
% identity
                  79
                  (AC004238) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7432
                   9123 1.R1040
Contig ID
5'-most EST
                  bth700845623.h1
                   7433
Seq. No.
Contig ID
                   9125 1.R1040
                  LIB3170-011-Q1-K1-B2
5'-most EST
Method
                  BLASTN
NCBI GI
                   g336391
                   79
BLAST score
                   3.0e-36
E value
Match length
                   159
% identity
                   87
NCBI Description A.thaliana chloroplast ribosomal protein S17 (rps17) mRNA,
                   3' end
Seq. No.
                   7434
Contig ID
                   9125 2.R1040
5'-most EST
                  LIB3049-024-Q1-E1-D12
Method
                   BLASTX
NCBI GI
                   q336392
BLAST score
                   148
E value
                   2.0e-09
Match length
                   53
% identity
                   58
                  (J05215) ribosomal protein S17 [Arabidopsis thaliana]
NCBI Description
```

7435

9128 1.R1040

Seq. No.

Contig ID



```
5'-most EST
                  uC-gmropic044e06b1
Seq. No.
                  7436
Contig ID
                   9132 1.R1040
5'-most EST
                   seb7\overline{0}0653953.h1
Method
                  BLASTX
NCBI GI
                  g3327868
BLAST score
                  140
E value
                   1.0e-08
Match length
                   89
                  35
% identity
                  (AB012912) COP1-Interacting Protein 7 (CIP7) [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  7437
Contig ID
                  9139 1.R1040
5'-most EST
                  LIB3039-007-Q1-E1-G9
Method
                  BLASTN
NCBI GI
                  q168492
BLAST score
                  229
                  1.0e-125
E value
Match length
                  385
% identity
                  90
NCBI Description Corn histone H3 (H3C3) gene, complete cds
Seq. No.
                  7438
                  9139 2.R1040
Contig ID
                  LIB3073-014-Q1-K1-D11
5'-most EST
Method
                  BLASTN
NCBI GI
                  g20252
BLAST score
                  179
                   6.0e-96
E value
Match length
                  371
                  87
% identity
NCBI Description Oryza sativa H3 histone H3R-21 clone RH3-2
Seq. No.
                  7439
                  9139 3.R1040
Contig ID
5'-most EST
                  LIB3049-049-Q1-E1-B3
Method
                  BLASTN
NCBI GI
                  q168492
BLAST score
                  173
E value
                  2.0e-92
Match length
                  297
% identity
                  90
NCBI Description Corn histone H3 (H3C3) gene, complete cds
Seq. No.
                  7440
                  9141 1.R1040
Contig ID
5'-most EST
                  LIB3051-048-Q1-K1-D3
Method
                  BLASTX
NCBI GI
                  g2598575
BLAST score
                  219
E value
                  2.0e-17
Match length
                  219
% identity
                  30
```

NCBI Description (Y15293) MtN21 [Medicago truncatula]



7441 Seq. No. 9148 1.R1040 Contig ID 5'-most EST LIB3039-027-Q1-E1-H3 Method BLASTX g2583134 NCBI GI BLAST score 143 6.0e-09 E value Match length 94 % identity 46 (AC002387) putative proline-rich protein [Arabidopsis NCBI Description thaliana] 7442 Seq. No. Contig ID 9154 1.R1040 LIB3040-035-Q1-E1-D7 5'-most EST 7443 Seq. No. 9154 2.R1040 Contig ID 5'-most EST LIB3107-050-Q1-K1-C1 Method BLASTN g940287 35

NCBI GI BLAST score E value 4.0e-10 Match length 75 87 % identity

NCBI Description Pisum sativum L. (clone na-481-5) mRNA, complete cds

7444 Seq. No. Contig ID 9155 1.R1040 5'-most EST LIB3039-027-Q1-E1-F6 Method BLASTX g3021598 NCBI GI BLAST score 171 2.0e-11 E value

Match length 346 % identity 25

(Y10389) nuclear protein [Xenopus laevis] NCBI Description

Seq. No. 7445

9155 2.R1040 Contig ID 5'-most EST leu701151007.hl

Seq. No. 7446

9159 1.R1040 Contig ID

5'-most EST LIB3087-007-Q1-K1-D5

7447 Seq. No.

9169 1.R1040 Contig ID

5'-most EST LIB3039-027-Q1-E1-E11

7448 Seq. No.

9171 1.R1040 Contig ID

5'-most EST LIB3039-027-Q1-E1-E5

BLASTX Method NCBI GI g3763934 BLAST score 226



E value 1.0e-18
Match length 124
% identity 50

NCBI Description (AC004450) hypothetical protein [Arabidopsis thaliana]

Seq. No.

7449

Contig ID

9176 1.R1040

5'-most EST

LIB3053-001-Q1-B1-A8

Method NCBI GI BLASTX g2384956

BLAST score E value g2384956 327

Match length % identity

3.0e-30 159 43

NCBI Description

(AF022985) No definition line found [Caenorhabditis

elegans]

Seq. No.

7450

Contig ID

9178 1.R1040

5'-most EST

fC-qmse7000762491d1

Seq. No.

7451

Contig ID

9179 1.R1040

5'-most EST

uC-gmropic029d10b1

Seq. No.

7452

Contig ID

9179 2.R1040

5'-most EST

LIB3049-023-Q1-E1-D11

Seq. No.

7453

Contig ID

9186_1.R1040

5'-most EST

 $jex7\overline{0}0909242.h1$

Method

BLASTX

NCBI GI

g585565

BLAST score

634

E value

8.0e-66

Match length

155

% identity

82

NCBI Description

RIBOSOME RECYCLING FACTOR HOMOLOG (NUCLEAR LOCATED PROTEIN

D2) >gi_629661_pir__S32716 nuclear protein - carrot

>gi_297891_emb_CAA51077_ (X72384) nuclear located protein

[Daucus carota]

Seq. No.

7454

Contig ID

9186_3.R1040

5'-most EST

epx701109783.h1

Seq. No.

7455

Contig ID

9194 1.R1040

5'-most EST

LIB3039-027-Q1-E1-A11

Seq. No.

7456

Contig ID

9203 1.R1040

5'-most EST

LIB3039-026-Q1-E1-H4

Seq. No.

7457

Contig ID

9212 1.R1040

5'-most EST g5688377 Method BLASTX NCBI GI g3879734 BLAST score 258 E value 8.0e-22 Match length 228 % identity 33

NCBI Description (Z93388) predicted using Genefinder; cDNA EST EMBL:D70912 comes from this gene; cDNA EST EMBL:D73452 comes from this gene; cDNA EST EMBL:D71574 comes from this gene; cDNA EST

EMBL:D74229 comes from this gene; cDNA EST EMBL:D727

7458 Seq. No.

Contig ID 9213 1.R1040 5'-most EST g5606317 Method BLASTX NCBI GI q4220482 BLAST score 423 2.0e-41 E value Match length 135 % identity 37

NCBI Description (AC006069) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7459

Contig ID 9213 2.R1040

5'-most EST LIB3039-026-Q1-E1-A1

Method BLASTX NCBI GI g4220482 BLAST score 226 E value 1.0e-18 Match length 74

% identity 59

NCBI Description (AC006069) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7460

Contig ID 9214 1.R1040

5'-most EST LIB3039-026-Q1-E1-F5

Method BLASTX NCBI GI g498040 BLAST score 555 E value 8.0e-57 Match length 191 % identity

NCBI Description (L33793) ORF [Senecio odorus]

Seq. No. 7461

9214 2.R1040 Contig ID

5'-most EST uC-gmropic102f04b1

Method BLASTX NCBI GI q498040 BLAST score 254 E value 6.0e-22 Match length 70 % identity 73

NCBI Description (L33793) ORF [Senecio odorus]

Seq. No. 7462



9219 1.R1040 Contig ID jC-gmst02400027f12a1 5'-most EST Method BLASTX NCBI GI g2511574 BLAST score 1117 E value 1.0e-122 Match length 235 % identity 94

NCBI Description (Y13176) multicatalytic endopeptidase [Arabidopsis thaliana] >gi_3421075 (AF043520) 20S proteasome subunit

PAB1 [Arabidopsis thaliana]

Seq. No. 7463

Contig ID 9219 2.R1040

5'-most EST LIB3049-052-Q1-E1-G3

Method BLASTX
NCBI GI g2511574
BLAST score 480
E value 2.0e-48
Match length 106
% identity 92

NCBI Description (Y13176) multicatalytic endopeptidase [Arabidopsis

thaliana] >gi_3421075 (AF043520) 20S proteasome subunit

PAB1 [Arabidopsis thaliana]

Seq. No. 7464

Contig ID 9222_1.R1040 5'-most EST vwf700676570.h1

Method BLASTX
NCBI GI g2981475
BLAST score 315
E value 5.0e-29
Match length 107
% identity 60

NCBI Description (AF053084) putative cinnamyl alcohol dehydrogenase [Malus

domestica]

Seq. No. 7465

Contig ID 9226_1.R1040 5'-most EST seb700654110.h1

Method BLASTX
NCBI GI g3219969
BLAST score 158
E value 3.0e-10
Match length 56
% identity 48

NCBI Description HYPOTHETICAL 22.4 KD PROTEIN C6G10.10C IN CHROMOSOME I

>gi 2330874 emb CAB11296 (Z98603) hypothetical protein

[Schizosaccharomyces pombe]

Seq. No. 7466

Contig ID 9227 1.R1040

5'-most EST LIB3 $\overline{0}$ 39-025-Q1-E1-F10

Method BLASTX
NCBI GI g1743354
BLAST score 593
E value 2.0e-61



Match length 179 % identity 60

NCBI Description (Y09876) aldehyde dehydrogenase (NAD+) [Nicotiana tabacum]

Seq. No. 7467

Contig ID 9228 1.R1040

5'-most EST LIB3039-026-Q1-E1-E5

Method BLASTN
NCBI GI g1778373
BLAST score 135
E value 1.0e-69
Match length 359
% identity 84

NCBI Description Pisum sativum glycine-rich RNA-binding protein PsGRBP mRNA,

complete cds

Seq. No. 7468

Contig ID 9239_1.R1040

5'-most EST LIB3039-026-Q1-E1-C7

Method BLASTX
NCBI GI g2500345
BLAST score 440
E value 3.0e-43
Match length 128
% identity 70

NCBI Description NHP2/RS6 FAMILY PROTEIN YEL026W HOMOLOG (HIGH MOBILITY

GROUP-LIKE NUCLEAR PROTEIN 2 HOMOLOG)

>gi_2618578_dbj_BAA23363_ (D50420) OTK27 [Homo sapiens]

>gi 3859990 (AF $\overline{0}$ 91076) OTK27 [Homo sapiens]

>gi 1589072 prf 2210268A nuclear protein-NHP2-like protein

[Homo sapiens]

Seq. No. 7469

Contig ID 9250_1.R1040 5'-most EST leu701150345.h1

Method BLASTX
NCBI GI g1477565
BLAST score 166
E value 2.0e-13
Match length 120
% identity 5

NCBI Description (U50078) p532 [Homo sapiens]

>gi 4557026 ref NP 003913.1 pHERC1 hect (homologous to the

E6-AP (UBE3A) carboxyl terminus) domain and RCC1

(CHC1)-like domain (RLD)

Seq. No. 7470

Contig ID 9251_1.R1040 5'-most EST smw700646267.h1

Method BLASTX
NCBI GI g2739168
BLAST score 1118
E value 1.0e-122
Match length 365
% identity 57

NCBI Description (AF032386) aldose-1-epimerase-like protein [Nicotiana

tabacum]

Method

BLASTX

```
7471
Seq. No.
Contig ID
                   9251 3.R1040
5'-most EST
                   jC-gmle01810047g07d1
                   7472
Seq. No.
Contig ID
                   9292 1.R1040
5'-most EST
                   LIB3039-024-Q1-E1-H5
Method
                   BLASTX
NCBI GI
                   g2829208
BLAST score
                   178
                   1.0e-12
E value
Match length
                   203
% identity
                   26
NCBI Description
                   (AF044076) candidate tumor suppressor p33ING1 [Homo
                   sapiens]
                   7473
Seq. No.
Contig ID
                   9295 1.R1040
5'-most EST
                  LIB3087-006-Q1-K1-B12
Method
                   BLASTX
NCBI GI
                   g2252866
BLAST score
                   230
E value
                   1.0e-18
Match length
                   69
% identity
                   65
NCBI Description
                   (AF013294) contains region of similarity to SYT
                   [Arabidopsis thaliana]
Seq. No.
                   7474
Contig ID
                   9295 2.R1040
5'-most EST
                   fC-gmst700653716f1
Method
                  BLASTX
NCBI GI
                   g2252866
BLAST score
                   177
                   1.0e-12
E value
Match length
                   48
% identity
                   69
NCBI Description
                   (AF013294) contains region of similarity to SYT
                   [Arabidopsis thaliana]
Seq. No.
                   7475
Contig ID
                   9295 3.R1040
5'-most EST
                   LIB3049-010-Q1-E1-B5
Method
                   BLASTX
NCBI GI
                   g2252866
BLAST score
                   194
E value
                   9.0e-15
Match length
                   49
% identity
                  76
NCBI Description
                   (AF013294) contains region of similarity to SYT
                   [Arabidopsis thaliana]
Seq. No.
                   7476
                  9295 4.R1040
Contig ID
5'-most EST
                  gsv701047345.h1
```

1377



```
g2252866
NCBI GI
                  190
BLAST score
E value
                  3.0e-14
Match length
                  52
% identity
                  69
                  (AF013294) contains region of similarity to SYT
NCBI Description
                   [Arabidopsis thaliana]
                  7477
Seq. No.
Contig ID
                  9295 5.R1040
                  crh700852275.h1
5'-most EST
Method
                  BLASTX
                  g2252866
NCBI GI
BLAST score
                  216
E value
                   4.0e-17
Match length
                  73
% identity
                   60
                   (AF013294) contains region of similarity to SYT
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                  7478
                   9295 6.R1040
Contig ID
                  fua701043038.hl
5'-most EST
                  7479
Seq. No.
Contig ID
                   9310 1.R1040
                   kl1701204787.h1
5'-most EST
                  7480
Seq. No.
                   9327 1.R1040
Contig ID
5'-most EST
                  LIB3039-023-Q1-E1-H11
                   7481
Seq. No.
                   9327 2.R1040
Contig ID
5'-most EST
                   LIB3040-048-Q1-E1-E11
                   7482
Seq. No.
Contig ID
                   9329 1.R1040
                   LIB3\overline{0}87-006-Q1-K1-G9
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3924823
                   255
BLAST score
                   1.0e-21
E value
Match length
                   135
% identity
                   38
                  (Z83113) cDNA EST yk491a11.5 comes from this gene; cDNA EST
NCBI Description
                   yk491a11.3 comes from this gene [Caenorhabditis elegans]
                   7483
Seq. No.
                   9329 2.R1040
Contig ID
5'-most EST
                   LIB3167-004-P1-K1-E7
                   BLASTX
Method
```

Method BLASTX
NCBI GI g3924823
BLAST score 203
E value 8.0e-16
Match length 71
% identity 51



NCBI Description (Z83113) cDNA EST yk491a11.5 comes from this gene; cDNA EST yk491a11.3 comes from this gene [Caenorhabditis elegans]

Seq. No. 7484

Contig ID 9345 1.R1040

5'-most EST LIB3039-023-Q1-E1-C7

Method BLASTN
NCBI GI g3021347
BLAST score 191
E value 1.0e-103
Match length 554
% identity 84

NCBI Description Cicer arietinum mRNA for cytoplasmic ribosomal protein L18

Seq. No. 7485

Contig ID 9345 3.R1040 q4284518 5'-most EST Method BLASTX g1172977 NCBI GI 270 BLAST score E value 1.0e-23 59 Match length 85 % identity

NCBI Description 60S RIBOSOMAL PROTEIN L18 >gi_606970 (U15741) cytoplasmic

ribosomal protein L18 [Arabidopsis thaliana]

Seq. No. 7486

Contig ID 9347_1.R1040 5'-most EST hrw701060933.h1

Seq. No. 7487

Contig ID 9355_1.R1040 5'-most EST uaw700666059.h1

Method BLASTX
NCBI GI g3775995
BLAST score 738
E value 2.0e-78
Match length 182
% identity 74

NCBI Description (AJ010461) RNA helicase [Arabidopsis thaliana]

Seq. No. 7488

Contig ID 9368 1.R1040

5'-most EST LIB3039-023-Q1-E1-A7

Method BLASTX
NCBI GI g2459438
BLAST score 276
E value 3.0e-24
Match length 74
% identity 66

NCBI Description (AC002332) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7489

Contig ID 9370 1.R1040

5'-most EST LIB3039-022-Q1-E1-G8

Seq. No. 7490



Contig ID 9382 1.R1040

5'-most EST uC-gmrominsoy310f07b1

Seq. No. 7491

Contig ID 9382 2.R1040

5'-most EST LIB3072-037-Q1-E1-H3

Seq. No. 7492

Contig ID 9385 1.R1040 5'-most EST g5607125 Method BLASTX g1723176 NCBI GI BLAST score 363 5.0e-34 E value 209 Match length % identity 40

NCBI Description HYPOTHETICAL 22.4 KD PROTEIN SLL0615

>gi_1001617_dbj_BAA10348_ (D64002) transmembrane protein
FT27 [Synechocystis sp.] >gi_1256592 (U38892) similar to
Mus musculus transmembrane protein (clone pFT27); Method:

conceptual translation supplied by author; ORF206

[Synechocystis sp.]

Seq. No. 7493

Contig ID 9385 2.R1040

5'-most EST LIB3106-110-Q1-K1-B12

Method BLASTX
NCBI GI g4539323
BLAST score 405
E value 6.0e-39
Match length 159
% identity 54

NCBI Description (AL035679) putative protein [Arabidopsis thaliana]

Seq. No. 7494

Contig ID 9398_1.R1040

5'-most EST LIB3039-022-Q1-E1-F3

Seq. No. 7495

Contig ID 9409_1.R1040

5'-most EST LIB3073-026-Q1-K1-G5

Method BLASTX
NCBI GI g3281868
BLAST score 197
E value 7.0e-15
Match length 73
% identity 59

NCBI Description (AL031004) putative protein [Arabidopsis thaliana]

Seq. No. 7496

Contig ID 9409 2.R1040

5'-most EST LIB3039-021-Q1-E1-H3

Seq. No. 7497

Contig ID 9411_1.R1040

5'-most EST LIB3073-011-Q1-K1-C10

Method BLASTX



NCBI GI g132825
BLAST score 215
E value 4.0e-17
Match length 80
% identity 62
NCBI Description 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR

>gi_71308_pir__R5PM25 ribosomal protein PsCL25 precursor,
chloroplast - garden pea >gi_20877_emb_CAA32187_ (X14022)

PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]

Seq. No. 7498

Contig ID 9411_2.R1040 5'-most EST pcp700988732.h1

Method BLASTN
NCBI GI g18729
BLAST score 73
E value 3.0e-33
Match length 93
% identity 95

NCBI Description Soybean (Glycine max) 18S ribosomal RNA

Seq. No. 7499

Contig ID 9418 1.R1040

5'-most EST % LIB3050-002-Q1-E1-G10

Method BLASTX
NCBI GI g2827554
BLAST score 167
E value 2.0e-11
Match length 53
% identity 62

NCBI Description (AL021635) putative DNA binding protein [Arabidopsis

thaliana]

Seq. No. 7500

Contig ID 9426_1.R1040 5'-most EST leu701157796.h1

Method BLASTX
NCBI GI g2997686
BLAST score 290
E value 1.0e-25
Match length 165
% identity 44

NCBI Description (AF053303) putative transcriptional co-activator

[Arabidopsis thaliana] >gi_3513735 (AF080118) contains similarity to RNA polymerase II transcription cofactor p15

[Arabidopsis thaliana] >gi_4539366_emb_CAB40060.1_(AL049525) putative protein [Arabidopsis thaliana]

Seq. No. 7501

Contig ID 9430_1.R1040 5'-most EST kll701213576.h1

Method BLASTX
NCBI GI g3482924
BLAST score 250
E value 3.0e-21
Match length 92
% identity 53



NCBI Description (AC003970) Highly similar to cinnamyl alcohol dehydrogenase, gi_1143445 [Arabidopsis thaliana]

Seq. No. 7502 Contig ID 9430_2.R1040

5'-most EST LIB3039-021-Q1-E1-F8

Method BLASTX
NCBI GI g3482926
BLAST score 323
E value 7.0e-30
Match length 118
% identity 55

NCBI Description (AC003970) Highly similar to cinnamyl alcohol dehydrogenase, gi 1143445 [Arabidopsis thaliana]

Seq. No. 7503

Contig ID 9433_1.R1040

5'-most EST LIB3049-021-Q1-E1-H2

Method BLASTX
NCBI GI g1173027
BLAST score 446
E value 5.0e-44
Match length 120
% identity 74

NCBI Description 60S RIBOSOMAL PROTEIN L31 >gi_915313 (U23784) ribosomal

protein L31 [Nicotiana glutinosa]

Seq. No. 7504

Contig ID 9435 1.R1040

5'-most EST LIB3039-021-Q1-E1-F1

Method BLASTX
NCBI GI g3132470
BLAST score 255
E value 6.0e-22
Match length 113
% identity 50

NCBI Description (AC003096) unknown protein [Arabidopsis thaliana]

Seq. No. 7505

Contig ID 9439_1.R1040 5'-most EST gsv701048747.h1

Seq. No. 7506

Contig ID 9446_1.R1040 5'-most EST zzp700835432.h1

Method BLASTX
NCBI GI g2147966
BLAST score 369
E value 5.0e-46
Match length 115
% identity 79

NCBI Description probable 1-acyl-sn-glycerol-3-phosphate acyltransferase -

Limnanthes douglasii >gi_1067138_emb_CAA88620_ (Z48730) 1-acyl-sn-glycerol-3-phosphate acyltransferase (putative)

[Limnanthes douglasii]

Seq. No. 7507

BLAST score

Match length

% identity

E value

188

214

32

9.0e-14



```
Contig ID
                  9451 1.R1040
5'-most EST
                  pxt700943324.h1
Method
                  BLASTX
                  g3152583
NCBI GI
BLAST score
                  204
E value
                  1.0e-15
Match length
                  209
% identity
                  34
NCBI Description
                  (AC002986) Contains similarity to inhibitor of apoptosis
                  protein gb_U45881 from D. melanogaster. [Arabidopsis
                  thalianal
Seq. No.
                  7508
                  9471 1.R1040
Contig ID
5'-most EST
                  LIB3106-050-Q1-K1-G2
Method
                  BLASTX
NCBI GI
                  g4514716
BLAST score
                  192
E value
                  2.0e-14
Match length
                  119
% identity
                  36
NCBI Description (AB017533) EPc [Nicotiana tabacum]
                  7509
Seq. No.
Contig ID
                  9478 1.R1040
5'-most EST
                  LIB3170-083-Q1-K1-E5
Method
                  BLASTX
NCBI GI
                  q3152565
BLAST score
                  412
                  7.0e-40
E value
Match length
                  251
% identity
                  44
                  (AC002986) EST gb N65759 comes from this gene. [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  7510
Contig ID
                  9482 1.R1040
5'-most EST
                  uC-gmrominsoy232d05b1
Seq. No.
                  7511
                  9482 2.R1040
Contig ID
5'-most EST
                  asn701141173.hl
Seq. No.
                  7512
                  9482 3.R1040
Contig ID
5'-most EST
                  LIB3107-074-Q1-K1-F11
Seq. No.
                  7513
                  9483 1.R1040
Contig ID
                  LIB3039-020-Q1-E1-H3
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2497226
```

NCBI Description HYPOTHETICAL 38.2 KD PROTEIN IN PRE5-FET4 INTERGENIC REGION

>gi_1363796_pir__S59308 probable membrane protein YMR315w yeast (Saccharomyces cerevisiae) >gi_984688_emb_CAA90833_
(Z54141) unknown [Saccharomyces cerevisiae]

Seq. No. 7514

Contig ID 9483_2.R1040

5'-most EST uC-gmropic067b04b1

Seq. No. 7515

Contig ID 9487_1.R1040 5'-most EST leu701145178.h1

Method BLASTX
NCBI GI g4115949
BLAST score 248
E value 3.0e-21
Match length 104
% identity 55

NCBI Description (AF118223) contains similarity to adenosine deaminases

[Arabidopsis thaliana]

Seq. No. 7516

Contig ID 9494_1.R1040 5'-most EST uxk700672041.h1

Method BLASTX
NCBI GI g1170504
BLAST score 1801
E value 0.0e+00
Match length 606
% identity 61

NCBI Description EUKARYOTIC INITIATION FACTOR (ISO)4F SUBUNIT P82

(IEIF-(ISO)4F P82) >gi 452440 (M95747) initiation factor

(iso)4f p82 subunit [Triticum aestivum]

Seq. No. 7517

Contig ID 9495_1.R1040 5'-most EST pcp700991891.h1

Seq. No. 7518

Contig ID 9511 1.R1040

5'-most EST LIB3074-004-Q1-K1-F12

Method BLASTX
NCBI GI g3724328
BLAST score 145
E value 1.0e-08
Match length 151
% identity 25

NCBI Description (AB013095) heme-binding protein [Mus musculus]

Seq. No. 7519

Contig ID 9511 2.R1040

5'-most EST LIB3109-003-Q1-K1-A6

Seq. No. 7520

Contig ID 9522 1.R1040 5'-most EST pcp700993234.h1

Method BLASTX NCBI GI g4314356



```
605
BLAST score
                  2.0e-62
E value
                  347
Match length
                  37
% identity
                  (AC006340) putative anthocyanidin-3-glucoside
NCBI Description
                  rhamnosyltransferase [Arabidopsis thaliana]
Seq. No.
                   9525 1.R1040
Contig ID
                  rlr700902209.h1
5'-most EST
                  BLASTX
Method
                   g3913633
NCBI GI
                   347
BLAST score
                   2.0e-32
E value
                   160
Match length
                   47
% identity
NCBI Description HYPOTHETICAL PROTEIN F8A5.25 >gi_2462742 (AC002292) Unknown
                   protein [Arabidopsis thaliana]
                   7522
Seq. No.
                   9529 1.R1040
Contig ID
                   qsv701055536.h1
5'-most EST
                   7523
Seq. No.
                   9532 1.R1040
Contig ID
                   LIB3109-031-Q1-K1-H7
5'-most EST
                   BLASTX
Method
                   q3894197
NCBI GI
                   152
BLAST score
                   7.0e-10
E value
                   82
Match length
                   38
% identity
                   (AC005662) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   7524
Seq. No.
                   9541 1.R1040
Contig ID
                   LIB3107-071-Q1-K1-A2
5'-most EST
                   BLASTX
Method
                   q2276463
NCBI GI
                   762
BLAST score
                   9.0e-81
E value
                   475
Match length
                   37
 % identity
                   (AF002020) Niemann-Pick C disease protein [Homo sapiens]
NCBI Description
                   >gi_4557803_ref_NP_000262.1_pNPC1_ Niemann-Pick disease,
                   type C1
                   7525
 Seq. No.
                   9547 1.R1040
 Contig ID
                   LIB3106-089-Q1-K1-F2
 5'-most EST
                   BLASTX
 Method
                   q2982432
 NCBI GI
                    325
 BLAST score
                    5.0e-30
 E value
                   73
 Match length
                   81
 % identity
                   (AL022224) putative protein [Arabidopsis thaliana]
 NCBI Description
```

7526 Seq. No. 9547 2.R1040 Contig ID LIB3106-067-P1-K1-F7 5'-most EST BLASTX Method NCBI GI g2982432 325 BLAST score 4.0e-30 E value 73 Match length 81 % identity NCBI Description (AL022224) putative protein [Arabidopsis thaliana] 7527 Seq. No. 9548 1.R1040 Contig ID leu701154367.hl 5'-most EST BLASTX Method NCBI GI g4544436 BLAST score 293 2.0e-26 E value 102 Match length 60 % identity (AC006592) anthocyanidin-3-glucoside rhamnosyltransferase, NCBI Description 3' partial [Arabidopsis thaliana] 7528 Seq. No. 9552 1.R1040 Contig ID kmv700740207.h1 5'-most EST BLASTX Method q2632106 NCBI GI BLAST score 711 3.0e-75E value 212 Match length % identity (Z98760) arginyl-tRNA synthetase [Arabidopsis thaliana] NCBI Description Seq. No. 7529 9558 1.R1040 Contig ID LIB3039-019-Q1-E1-G11 5'-most EST 7530 Seq. No. 9558 2.R1040 Contig ID jC-gmf102220075f08a1 5'-most EST 7531 Seq. No. 9560 1.R1040 Contig ID vzy700754495.hl 5'-most EST 7532 Seq. No. 9562 1.R1040 Contig ID

leu701157074.h1 5'-most EST

7533 Seq. No.

9565 1.R1040 Contig ID

LIB3170-026-Q1-K1-C10 5'-most EST

BLASTX Method g3876465 NCBI GI BLAST score 329



2.0e-30 E value 83 Match length % identity (Z81071) predicted using Genefinder; Similarity to Human NCBI Description small nuclear ribonucleoprotein E (SW:P08578); cDNA EST yk375g7.3 comes from this gene; cDNA EST yk375g7.5 comes from this gene; cDNA EST yk435f5.3 comes from this gene; 7534 Seq. No. 9568 1.R1040 Contig ID wvk700686485.hl 5'-most EST BLASTX Method g3668074 NCBI GI 526 BLAST score 2.0e-53 E value 159 Match length % identity NCBI Description (AC004667) hypothetical protein [Arabidopsis thaliana] Seq. No. 9585 1.R1040 Contig ID txt700734422.h1 5'-most EST BLASTX Method q3482973 NCBI GI 293 BLAST score 3.0e-26 E value 168 Match length % identity NCBI Description (AL031369) putative protein [Arabidopsis thaliana] 7536 Seq. No. 9585 2.R1040 Contig ID 5'-most EST LIB3039-019-Q1-E1-D7 Method BLASTX q3482973 NCBI GI 144 BLAST score 8.0e-09 E value Match length 98 % identity (AL031369) putative protein [Arabidopsis thaliana] NCBI Description 7537 Seq. No. Contig ID 9593 1.R1040 ncj700982214.h1 5'-most EST 7538 Seq. No. 9593 2.R1040 Contig ID kl1701204547.h2 5'-most EST 7539 Seq. No. 9597 1.R1040 Contig ID $LIB3\overline{0}39-012-Q1-E1-A5$

5'-most EST

Method BLASTX g3738309 NCBI GI

303 BLAST score 2.0e-27 E value 111 Match length



```
% identity
                 (AC005309) unknown protein [Arabidopsis thaliana]
NCBI Description
                  7540
Seq. No.
                  9600 1.R1040
Contig ID
                  leu701150153.hl
5'-most EST
                  BLASTX
Method
                  g549577
NCBI GI
                  227
BLAST score
                  1.0e-18
E value
                  66
Match length
                  65
% identity
                  HYPOTHETICAL 17.1 KD PROTEIN IN MECB-GLTX INTERGENIC REGION
NCBI Description
                  >gi_2127059_pir_ S66120 hypothetical protein - Bacillus
                  subtilis >gi 289281 (L14580) unknown [Bacillus subtilis]
                  >gi_467479_dbj_BAA05325_ (D26185) unknown [Bacillus
                  subtilis] >gi 2632358_emb_CAB11867_ (Z99104) similar to
                  hypothetical proteins [Bacillus subtilis]
                  7541
Seq. No.
                  9601 1.R1040
Contig ID
                  epx701107762.h1
5'-most EST
                  7542
Seq. No.
                  9603 1.R1040
Contig ID
                  jC-qmle01810089e12d1
5'-most EST
                  7543
Seq. No.
                  9606 1.R1040
Contig ID
                  leu701152657.hl
5'-most EST
                  BLASTX
Method
                  q4455202
NCBI GI
BLAST score
                   405
                   2.0e-39
E value
                  114
Match length
% identity
NCBI Description (AL035440) putative APG protein [Arabidopsis thaliana]
                   7544
Seq. No.
                   9610 1.R1040
Contig ID
                   rlr700901982.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4559310
BLAST score
                   154
                   1.0e-09
E value
                   119
Match length
% identity
                   (AF129131) putative Zic3 binding protein; CBP3 protein
NCBI Description
                   homolog [Xenopus laevis]
```

Seq. No. 7545

Contig ID 9626_1.R1040 5'-most EST gsv701051404.h1

Method BLASTX
NCBI GI g2194122
BLAST score 221
E value 1.0e-17

Match length 151 % identity 40

NCBI Description (AC002062) No definition line found [Arabidopsis thaliana]

Seq. No. 7546

Contig ID 9626_2.R1040

5'-most EST LIB3051-020-Q1-E1-G3

Method BLASTX
NCBI GI g3123176
BLAST score 287
E value 4.0e-25
Match length 249
% identity 22

NCBI Description HYPOTHETICAL 43.1 KD TRP-ASP REPEATS CONTAINING PROTEIN

KO4G11.4 IN CHROMOSOME X >gi_3878300_emb_CAB01760_ (Z78544) predicted using Genefinder; Similarity to C.elegans Guanine nucleotide binding protein (WP:C14B1.4) [Caenorhabditis

elegans]

Seq. No. 7547

Contig ID 9626_3.R1040

5'-most EST LIB3039-018-Q1-E1-E3

Method BLASTX
NCBI GI g3451472
BLAST score 155
E value 1.0e-09
Match length 112
% identity 9

NCBI Description (AL031349) transcription initiation factor TFIID subunit

[Schizosaccharomyces pombe]

Seq. No. 7548

Contig ID 9626_6.R1040

5'-most EST jC-gmle01810035d06d1

Seq. No. 7549

Contig ID 9630_1.R1040

5'-most EST jC-gmf102220114c12a1

Method BLASTX
NCBI GI g3367537
BLAST score 372
E value 4.0e-35
Match length 137
% identity 54

NCBI Description (AC004392) Contains similarity to ANK repeat region of

Fowlpox virus BamHI-orf7 protein homolog C18F10.7 gi_485107 from Caenorhabditis elegans cosmid gb_U00049. This gene is continued from unannotated gene on BAC F19K23 gb_AC000375.

[Arabid

Seq. No. 7550

Contig ID 9638_1.R1040

5'-most EST LIB3039-018-Q1-E1-D1

Seq. No. 7551

Contig ID 9644 1.R1040

5'-most EST jC-gmf102220053a07a1



```
7552
Seq. No.
                  9646 1.R1040
Contig ID
5'-most EST
                  LIB3039-018-Q1-E1-C2
Method
                  BLASTX
NCBI GI
                  g3451067
BLAST score
                  189
E value
                  2.0e-14
Match length
                  44
                  84
% identity
NCBI Description
                 (AL031326) putative protein [Arabidopsis thaliana]
                  7553
Seq. No.
                  9665 1.R1040
Contig ID
5'-most EST
                  LIB3039-017-Q1-E1-H4
Seq. No.
                  7554
                  9666 1.R1040
Contig ID
5'-most EST
                  LIB3039-017-Q1-E1-H5
Seq. No.
                  7555
Contig ID
                  9672 1.R1040
5'-most EST
                  LIB3039-018-Q1-E1-A12
Method
                  BLASTX
NCBI GI
                  g632220
BLAST score
                  315
                  6.0e-29
E value
                  112
Match length
% identity
                  61
                  ribosomal protein L2 - evening primrose mitochondrion
NCBI Description
                  >gi 516394 emb CAA56451 (X80170) 70s mitochondrial
                  ribosomal protein L2 [Oenothera berteriana]
Seq. No.
                  7556
Contig ID
                  9673 1.R1040
5'-most EST
                  LIB3039-018-Q1-E1-A3
Method
                  BLASTX
NCBI GI
                  q2253579
BLAST score
                  141
                  9.0e-09
E value
                  83
Match length
                  45
% identity
                  (U78721) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  7557
                  9676 1.R1040
Contig ID
5'-most EST
                  LIB3039-017-Q1-E1-E3
Method
                  BLASTX
NCBI GI
                  g109219
BLAST score
                  272
E value
                  1.0e-23
Match length
                  230
% identity
                  34
```

Seq. No. 7558

NCBI Description dimethylaniline monooxygenase (N-oxide-forming) (EC 1.14.13.8), hepatic 1 - rabbit

```
9677 1.R1040
Contig ID
                  q5677150
5'-most EST
                  7559
Seq. No.
                  9680 1.R1040
Contig ID
5'-most EST
                  bth700846387.h1
                  7560
Seq. No.
                  9690 1.R1040
Contig ID
                  jC-gmle01810086e11a1
5'-most EST
                  7561
Seq. No.
                  9702 1.R1040
Contig ID
                  LIB3039-017-Q1-E1-C11
5'-most EST
                  BLASTX
Method
                  g2262165
NCBI GI
BLAST score
                  255
                   4.0e-22
E value
Match length
                   101
                   51
% identity
NCBI Description (AC002329) predicted protein of unknown function
                   [Arabidopsis thaliana]
                   7562
Seq. No.
                   9717 1.R1040
Contig ID
                   LIB3039-016-Q1-E1-G5
5'-most EST
                   BLASTX
Method
                   q3097321
NCBI GI
                   550
BLAST score
                   3.0e-58
E value
                   283
Match length
% identity
NCBI Description (AB013289) Bd 30K [Glycine max]
                   7563
Seq. No.
                   9722 1.R1040
Contig ID
                   LIB3039-016-Q1-E1-H11
5'-most EST
                   BLASTX
Method
                   g1699370
NCBI GI
                   296
BLAST score
                   1.0e-26
E value
Match length
                   149
% identity
                   46
                   (S82620) prolidase=peptidase D/imidopeptidase {EC 3.4.13.9}
NCBI Description
                   [mice, liver, Peptide, 493 aa] [Mus sp.]
                   7564
Seq. No.
                   9727_1.R1040
Contig ID
                   epx701107982.h1
5'-most EST
                   BLASTX
Method
                   g3415134
NCBI GI
BLAST score
                   165
                   2.0e-11
E value
                   40
Match length
                   75
 % identity
```

NCBI Description (AF082024) Phybl [Pimpinella brachycarpa]



```
7565
Seq. No.
                  9727 2.R1040
Contig ID
                  LIB3139-041-P1-N1-H5
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3415134
BLAST score
                  279
                  6.0e-25
E value
                  76
Match length
% identity
NCBI Description (AF082024) Phyb1 [Pimpinella brachycarpa]
Seq. No.
                  9728 1.R1040
Contig ID
                  ncj700981368.hl
5'-most EST
Seq. No.
                  9728 2.R1040
Contig ID
5'-most EST
                  wrq700790469.h2
                  7568
Seq. No.
                  9732 1.R1040
Contig ID
5'-most EST
                  LIB3049-054-Q1-E1-H2
                  BLASTX
Method
                  g4056506
NCBI GI
                  201
BLAST score
                   2.0e-15
E value
Match length
                   137
% identity
NCBI Description (AC005896) nodulin-like protein [Arabidopsis thaliana]
                   7569
Seq. No.
Contig ID
                   9732 2.R1040
                   LIB3092-058-Q1-K1-E1
5'-most EST
                   7570
Seq. No.
                   9732 3.R1040
Contig ID
5'-most EST
                   zhf700960794.h1
Method
                   BLASTX
                   g2623304
NCBI GI
BLAST score
                   204
                   7.0e-16
E value
Match length
                   132
% identity
                   (ACO02409) similar to Medicago nodulin N21 [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7571
                   9733 1.R1040
Contig ID
                   LIB3040-056-Q1-E1-D10
5'-most EST
                   7572
Seq. No.
Contig ID
                   9736 1.R1040
                   LIB3039-016-Q1-E1-F6
5'-most EST
```

BLASTX Method g3413703 NCBI GI BLAST score 263 1.0e-22 E value



```
184
Match length
                   45
% identity
                   (AC004747) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                    9739 1.R1040
                   pcp700990081.hl
5'-most EST
                    7574
Seq. No.
                    9744 2.R1040
Contig ID
5'-most EST
                    gsv701056547.h1
                    7575
Seq. No.
                    9744 3.R1040
Contig ID
                    crh700856518.h1
5'-most EST
                    7576
Seq. No.
                    9751 1.R1040
Contig ID
                    pmv700895050.hl
5'-most EST
                    7577
Seq. No.
                    9755 1.R1040
Contig ID
                    hrw701061761.hl
5'-most EST
                    BLASTX
Method
                    g4454051
NCBI GI
                    267
BLAST score
                    8.0e-35
E value
                    147
Match length
% identity
                    (AL035394) putative polygalacturonase [Arabidopsis
NCBI Description
                    thaliana]
Seq. No.
                    7578
                    9757 1.R1040
Contig ID
                    LIB3039-016-Q1-E1-C4
5'-most EST
Seq. No.
                    7579
                    9769 1.R1040
Contig ID
                    fde700875709.h1
 5'-most EST
                    BLASTX
Method
                    q3600032
 NCBI GI
                    240
 BLAST score
                     4.0e-20
 E value
                    128
 Match length
                     48
 % identity
                    (AF080119) contains similarity to tropomyosin (Pfam: Tropomyosin.hmm, score: 14.57) and ATP synthase (Pfam:
 NCBI Description
                    ATP-synt_B.hmm, score: 10.89) [Arabidopsis thaliana]
                     7580
 Seq. No.
```

Contig ID 9770_1.R1040

5'-most EST LIB3040-003-Q1-E1-H1

Method BLASTX
NCBI GI g3687251
BLAST score 546
E value 9.0e-56
Match length 143



% identity (AC005169) unknown protein [Arabidopsis thaliana] NCBI Description 7581 Seq. No. 9784 1.R1040 Contig ID LIB3107-003-Q1-K1-B10 5'-most EST Seq. No. 7582 9785 1.R1040 Contig ID $xpa7\overline{0}0795562.h1$ 5'-most EST 7583 Seq. No. 9788 1.R1040 Contig ID q5606786 5'-most EST BLASTX Method g4539390 NCBI GI 1815 BLAST score 0.0e+00E value 380 Match length 88 % identity (AL035526) shaggy-like protein kinase etha (EC 2.7.1.-) NCBI Description [Arabidopsis thaliana] 7584 Seq. No. 9788 2.R1040 Contig ID jC-gmst02400055c01a1 5'-most EST BLASTX Method g4056456 NCBI GI 686 BLAST score 3.0e-72 E value Match length 204 % identity (AC005990) Strong similarity to gb_U20808 auxin-induced NCBI Description protein from Vigna radiata and a member of the zinc-binding dehydrogenase family PF_00107. ESTs gb_T43674, gb_H77006 and gb_AA395179 come from this gene. [Arabidopsis thaliana] Seq. No. 9788 3.R1040 Contig ID crh700851255.hl 5'-most EST BLASTX Method NCBI GI q4539390 BLAST score 626 E value 4.0e-65 135 Match length 87 % identity (AL035526) shaggy-like protein kinase etha (EC 2.7.1.-) NCBI Description [Arabidopsis thaliana] 7586 Seq. No. 9788_4.R1040 Contig ID LIB3049-020-Q1-E1-B6 5'-most EST Method BLASTX

1394

g2182029

1.0e-61

579

148

NCBI GI

E value

BLAST score

Match length

% identity NCBI Description (Y13437) shaggy-like kinase etha (OSKetha) [Oryza sativa] 7587 Seq. No. 9788 5.R1040 Contig ID xpa700792520.hl 5'-most EST BLASTX Method g2129738 NCBI GI 560 BLAST score 1.0e-57 E value 141 Match length 74 % identity NCBI Description shaggy-like kinase dzeta - Arabidopsis thaliana >gi 1225913 emb_CAA64408_ (X94938) shaggy-like kinase dzeta [Arabidopsis thaliana] >gi_1669653_emb_CAA70483_ (Y09300) serine/threonine kinase [Arabidopsis thaliana] Seq. No. 9788 6.R1040 Contig ID uC-gmflminsoy079c09b1 5'-most EST Seq. No. 9788 9.R1040 Contig ID LIB3170-024-Q1-J1-A12 5'-most EST 7590 Seq. No. 9788 11.R1040 Contig ID LIB3049-002-Q1-E1-G7 5'-most EST BLASTX Method g1532168 NCBI GI BLAST score 194 7.0e-15 E value 66 Match length % identity (U63815) localized according to blastn similarity to EST NCBI Description sequences; therefore, the coding span corresponds only to an area of similarity since the initation codon and stop codon could not be precisely determined [Arabidopsis thaliana] 7591 Seq. No. 9788_12.R1040 Contig ID uC-qmflminsoy069b05b1 5'-most EST Method BLASTX q4539390 NCBI GI 413 BLAST score 2.0e-52 E value 119 Match length % identity 81 (AL035526) shaggy-like protein kinase etha (EC 2.7.1.-)

NCBI Description

[Arabidopsis thaliana]

Seq. No. 7592

9788 13.R1040 Contig ID

LIB3049-020-Q1-E1-B4 5'-most EST

BLASTX Method g1877397 NCBI GI



```
207
BLAST score
E value
                  4.0e-16
Match length
                  48
% identity
                  81
NCBI Description (Y11591) shaggy-like kinase [Ricinus communis]
                  7593
Seq. No.
                  9792 1.R1040
Contig ID
5'-most EST
                  LIB3039-015-Q1-E1-G10
Method
                  BLASTX
NCBI GI
                  g4335754
BLAST score
                  191
E value
                  2.0e-23
                  76
Match length
                  71
% identity
NCBI Description
                  (AC006284) putative acid phophatase (contains
                  metallo-phosphoesterase motif, prosite: QDOC50185)
                  [Arabidopsis thaliana]
                  7594
Seq. No.
                  9799 1.R1040
Contig ID
5'-most EST
                  zsg701123467.h1
Method
                  BLASTX
NCBI GI
                  g3702332
BLAST score
                  480
E value
                  8.0e-48
Match length
                  222
% identity
                  42
NCBI Description
                  (AC005397) unknown protein [Arabidopsis thaliana]
                  7595
Seq. No.
Contig ID
                  9799 2.R1040
5'-most EST
                  LIB3106-085-Q1-K1-E12
Method
                  BLASTX
                  q3702332
NCBI GI
BLAST score
                  243
E value
                  1.0e-20
Match length
                  82
% identity
NCBI Description
                  (AC005397) unknown protein [Arabidopsis thaliana]
Seq. No.
                  7596
Contig ID
                  9800 1.R1040
5'-most EST
                  LIB3039-015-Q1-E1-D9
                  7597
Seq. No.
                  9800 2.R1040
Contig ID
5'-most EST
                  zsg701129547.h1
```

7598 Seq. No.

9804 1.R1040 Contig ID

5'-most EST LIB3056-013-Q1-N1-B12

Seq. No. 7599

9807 1.R1040 Contig ID 5'-most EST vzy700750953.h1

BLASTX Method

NCBI GI g2854153 BLAST score 296 E value 1.0e-35 Match length 166 % identity 51

NCBI Description (AF045640) No definition line found [Caenorhabditis

elegans]

Seq. No. 7600

Contig ID 9807 2.R1040

5'-most EST LIB3093-039-Q1-K1-A3

Method BLASTX
NCBI GI g2854153
BLAST score 149
E value 2.0e-09
Match length 55
% identity 56

NCBI Description (AF045640) No definition line found [Caenorhabditis

elegans]

Seq. No. 7601

Contig ID 9811_1.R1040 5'-most EST leu701154049.h1

Seq. No. 7602

Contig ID 9813 1.R1040

5'-most EST LIB3039-015-Q1-E1-D8

Seq. No. 7603

Contig ID 9815_1.R1040 5'-most EST trc700562769.h1

Method BLASTX
NCBI GI g3063710
BLAST score 188
E value 2.0e-13
Match length 116
% identity 34

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 7604

Contig ID 9815 4.R1040

5'-most EST jC-gmro02910019e01a1

Seq. No. 7605

Contig ID 9818_1.R1040

5'-most EST LIB3039-015-Q1-E1-C7

Method BLASTX
NCBI GI g3892050
BLAST score 550
E value 1.0e-81
Match length 281
% identity 67

NCBI Description (AC002330) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7606

Contig ID 9830 1.R1040

5'-most EST LIB3039-015-Q1-E1-B10

```
7607
Seq. No.
                   9830 2.R1040
Contig ID
                  LIB3106-095-Q1-K1-D6
5'-most EST
                  7608
Seq. No.
                   9830 3.R1040
Contig ID
                  uC-gmflminsoy017c12b1
5'-most EST
Seq. No.
                   9841 1.R1040
Contig ID
                   hrw701062003.h1
5'-most EST
                   7610
Seq. No.
                   9858 1.R1040
Contig ID
                   xpa700795673.h1
5'-most EST
                   BLASTX
Method
                   q1532175
NCBI GI
BLAST score
                   408
                   1.0e-39
E value
                   104
Match length
                   68
% identity
                   (U63815) similar to protein disulfide isomerase
NCBI Description
                   [Arabidopsis thaliana]
                   7611
Seq. No.
                   9859 1.R1040
Contig ID
                   LIB3039-001-Q1-E1-H4
5'-most EST
                   BLASTX
Method
                   q3128177
NCBI GI
                   508
BLAST score
                   3.0e-51
E value
Match length
                   136
% identity
                   (AC004521) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7612
Seq. No.
                   9859 2.R1040
Contig ID
                   pcp700991855.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3128177
                   294
BLAST score
                   1.0e-26
E value
                   79
Match length
 % identity
                   73
                    (AC004521) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7613
                    9862 1.R1040
 Contig ID
                    gsv7\overline{0}1048517.h1
 5'-most EST
                   BLASTX
Method
                    g3236245
 NCBI GI
                    635
 BLAST score
                    7.0e-72
 E value
                    241
 Match length
                    63
 % identity
                   (AC004684) myb-related protein [Arabidopsis thaliana]
 NCBI Description
```

1398

```
7614
Seq. No.
                    9862 2.R1040
Contig ID
                    6HA - \overline{0}2 - Q1 - B1 - B10
5'-most EST
                    BLASTX
Method
NCBI GI
                    q3236245
                    333
BLAST score
                    5.0e-31
E value
                    73
Match length
% identity
NCBI Description (AC004684) myb-related protein [Arabidopsis thaliana]
                    7615
Seq. No.
                    9865 1.R1040
Contig ID
                    hrw701061053.h1
5'-most EST
                    BLASTX
Method
NCBI GI
                    q1171870
                    163
BLAST score
                    7.0e-11
E value
                    93
Match length
                    33
% identity
                    NADH-UBIQUINONE OXIDOREDUCTASE 19 KD SUBUNIT (COMPLEX
NCBI Description
                    I-19KD) (CI-19KD) (COMPLEX I-PGIV) (CI-PGIV)
                    >gi 89670 pir_ S16208 NADH dehydrogenase (ubiquinone) (EC
                    1.6.5.3) 19K chain - bovine >gi_599681_emb_CAA42218
                     (X59697) 19 kDa subunit of NADH: ubiquinone oxidoreductase
                    complex (complex I) [Bos taurus]
                    7616
Seq. No.
                     9867 1.R1040
Contig ID
5'-most EST
                    LIB3039-014-Q1-E1-E6
                    BLASTX
Method
NCBI GI
                     q1351018
                     391
BLAST score
                     1.0e-37
E value
                     99
Match length
% identity
                     RSI-1 PROTEIN PRECURSOR (TR132) >gi_1076597_pir__S43909
NCBI Description
                     RSI-1 protein - tomato >gi_405585 (L22188) RSI-1 protein
                     [Lycopersicon esculentum] \overline{>}gi_405587 (L22189) RSI-1 protein
                     [Lycopersicon esculentum]
                     7617
 Seq. No.
                     9872 1.R1040
 Contig ID
                     LIB3072-018-Q1-E1-G1
 5'-most EST
                     BLASTX
Method
                     q1172995
NCBI GI
 BLAST score
                     251
                     3.0e-21
 E value
                     99
 Match length
                     57
 % identity
                     60S RIBOSOMAL PROTEIN L22 >gi_1083790_pir__S52084 ribosomal protein L22 - rat >gi_710295_emb_CAA55204_ (X78444) ribosomal protein L22 [Rattus norvegicus]
 NCBI Description
                     >gi_1093952_prf__2105193A ribosomal protein L22 [Rattus
```

norvegicus]



```
7618
Seq. No.
                  9883 1.R1040
Contig ID
                  LIB3039-014-Q1-E1-D1
5'-most EST
                  BLASTX
Method
                  g1076531
NCBI GI
BLAST score
                  728
                   3.0e-77
E value
                  152
Match length
% identity
                  hypothetical protein, pollen allergen homolog - garden pea
NCBI Description
                   >gi 2129891_pir__S65056 pollen allergen homolog precursor
                   (clone PPA1) - garden pea >gi_732905_emb_CAA59470_ (X85187)
                   homology with pollen allergens [Pisum sativum]
                   7619
Seq. No.
                   9886 1.R1040
Contig ID
5'-most EST
                   LIB3039-014-Q1-E1-A12
                   BLASTX
Method
                   q4249414
NCBI GI
                   245
BLAST score
                   1.0e-20
E value
                   113
Match length
% identity
NCBI Description (AC006072) unknown protein [Arabidopsis thaliana]
Seq. No.
                   9886 3.R1040
Contig ID
                   ncj700979452.hl
5'-most EST
                   BLASTN
Method
                   q4249393
NCBI GI
                   41
BLAST score
                   7.0e-14
E value
Match length
                   81
% identity
                   Arabidopsis thaliana chromosome II BAC T9J23 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
                   7621
 Seq. No.
                   9890 1.R1040
 Contig ID
                   LIB3073-005-Q1-K1-F1
 5'-most EST
 Seq. No.
                   7622
                   9908 1.R1040
 Contig ID
                   fua701041841.h1
 5'-most EST
                   BLASTX
 Method
                   g3915186
 NCBI GI
                   344
 BLAST score
                   5.0e - 32
 E value
                   147
 Match length
                    47
 % identity
                   UBIQUITIN-CONJUGATING ENZYME E2-21 KD (UBIQUITIN-PROTEIN
```

7623 Seq. No.

NCBI Description

9908 2.R1040 Contig ID

[Pichia angusta]

LIGASE) (UBIQUITIN CARRIER PROTEIN) (PEROXIN-4) >gi_3128447 (AF061604) ubiquitin-conjugating enzyme homolog peroxin 4

5'-most EST

```
5'-most EST
                   uC-gmflminsoy037b04b1
Method
                   BLASTX
NCBI GI
                   g3273562
BLAST score
                   255
E value
                   8.0e-22
Match length
                   104
% identity
                   45
NCBI Description
                  (AF036707) RAD6 [Candida albicans]
Seq. No.
                   7624
Contig ID
                   9915 1.R1040
5'-most EST
                  LIB3039-013-Q1-E1-F6
Method
                  BLASTX
NCBI GI
                  g285741
BLAST score
                   152
                   6.0e-10
E value
Match length
                  79
% identity
                  39
NCBI Description
                  (D14550) EDGP precursor [Daucus carota]
                  7625
Seq. No.
Contig ID
                  9927 1.R1040
5'-most EST
                  LIB3107-040-Q1-K1-C1
Method
                  BLASTX
NCBI GI
                  g3024629
BLAST score
                  201
E value
                  2.0e-15
Match length
                  137
% identity
                  33
NCBI Description
                  GAMMA-SOLUBLE NSF ATTACHMENT PROTEIN (SNAP-GAMMA)
                  >gi_423252_pir__S32369 gamma-SNAP protein - bovine
                  >gi 298669 bbs 127528 gamma soluble NSF attachment protein.
                  gamma SNAP=N-ethyl-maleimide-sensitive fusion protein
                  attachment protein [cattle, brain, Peptide, 328 aa]
                  >gi 445846 prf 1910317C NSF attachment protein
                   (SNAP):ISOTYPE=gamma [Bos taurus]
Seq. No.
                  7626
Contig ID
                  9930 1.R1040
5'-most EST
                  LIB3139-107-P1-N1-B3
Method
                  BLASTX
NCBI GI
                  q2494113
BLAST score
                  1038
E value
                  1.0e-113
Match length
                  226
% identity
NCBI Description
                  (AC002376) Strong similarity to Musa pectate lyase
                  (gb_X92943). ESTs gb_AA042458, gb_ATTS4502, gb_N38552 come
                  from this gene. [Arabidopsis thaliana]
Seq. No.
                  7627
Contig ID
                  9933 1.R1040
5'-most EST
                  LIB3052-012-Q1-N1-B3
Seq. No.
                  7628
                  9938 1.R1040
Contig ID
```

1401

sat701015315.h1

BLAST score

Match length

% identity

E value

647

138

89

1.0e-67



```
Method
                   BLASTX
NCBI GI
                   q4538965
BLAST score
                   250
E value
                   2.0e-21
Match length
                  105
% identity
                   47
NCBI Description
                  (AL049488) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  7629
Contig ID
                  9942 1.R1040
5'-most EST
                  LIB3170-055-Q1-K1-G2
Method
                  BLASTX
NCBI GI
                  g2494162
BLAST score
                  776
E value
                  2.0e-82
Match length
                  335
% identity
                  49
NCBI Description
                  HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II
                  PRECURSOR >gi 3879803 emb CAA87414 (Z47356) similar to
                  DNAJ [Caenorhabditis elegans]
Seq. No.
                  7630
Contig ID
                  9942 4.R1040
5'-most EST
                  uC-gmrominsoy128f10b1
Method
                  BLASTX
NCBI GI
                  g144832
BLAST score
                  152
E value
                  6.0e-10
Match length
                  41
% identity
                  63
NCBI Description
                  (M74569) dnaJ [Clostridium acetobutylicum]
                  7631
Seq. No.
Contig ID
                  9944 1.R1040
5'-most EST
                  LIB3039-013-Q1-E1-B3
Seq. No.
                  7632
Contig ID
                  9944 2.R1040
5'-most EST
                  uC-gmflminsoy056e05b1
Seq. No.
                  7633
Contig ID
                  9948 1.R1040
5'-most EST
                  uC-gmrominsoy271a10b1
Seq. No.
                  7634
Contig ID
                  9948 2.R1040
5'-most EST
                  LIB3109-014-Q1-K1-B8
Seq. No.
                  7635
Contig ID
                  9949 1.R1040
5'-most EST
                  jC-gmro02800041h12a1
Method
                  BLASTX
NCBI GI
                  q2760362
```

Contig ID

7641

9956 3.R1040



```
(AF016511) 15.9 kDa subunit of RNA polymerase II
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   7636
Contig ID
                   9949 2.R1040
5'-most EST
                   xpa700797793.h1
Method
                   BLASTX
NCBI GI
                   q2760362
BLAST score
                   633
E value
                   6.0e-66
Match length
                   138
% identity
                   88
NCBI Description
                   (AF016511) 15.9 kDa subunit of RNA polymerase II
                   [Arabidopsis thaliana]
Seq. No.
                   7637
Contig ID
                   9949 4.R1040
5'-most EST
                   rca700999513.h1
Method
                   BLASTN
                   g2760361
NCBI GI
BLAST score
                   120
E value
                   7.0e-61
Match length
                   240
% identity
                   88
NCBI Description
                  Arabidopsis thaliana 15.9 kDa subunit of RNA polymerase II
                   (RPB15.9) mRNA, complete cds
Seq. No.
                   7638
Contig ID
                   9955 1.R1040
5'-most EST
                   jC-gmle01810008e01a1
Method
                   BLASTX
NCBI GI
                   g3928543
BLAST score
                   1499
E value
                   1.0e-167
Match length
                   475
% identity
                   58
                   (AB016819) UDP-glucose glucosyltransferase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7639
Contig ID
                   9956 1.R1040
5'-most EST
                   jC-gmro02910018a04a1
Seq. No.
                   7640
                   9956 2.R1040
Contig ID
5'-most EST
                  LIB3039-012-Q1-E1-H8
Method
                  BLASTX
NCBI GI
                   g4309884
BLAST score
                   284
E value
                   2.0e-25
Match length
                  79
% identity
                  65
NCBI Description
                   (AC006389) similar to Schizosaccharomyces pombe splicing
                  factor; similar to PID:3395591 [Homo sapiens]
```

1403

```
5'-most EST
                   LIB3074-026-Q1-E1-C4
Method
                   BLASTX
NCBI GI
                   g4309884
BLAST score
                   116
E value
                   2.0e-09
Match length
                   68
% identity
                   53
NCBI Description
                  (AC006389) similar to Schizosaccharomyces pombe splicing
                   factor; similar to PID:3395591 [Homo sapiens]
                   7642
Seq. No.
Contig ID
                   9958 1.R1040
5'-most EST
                   seb7\overline{0}0649775.h1
Method
                   BLASTN
NCBI GI
                   q410285
BLAST score
                   367
E value
                   0.0e + 00
Match length
                   804
% identity
                   90
NCBI Description
                  Pisum sativum rho (ras-related) GTP-binding protein mRNA,
                   complete cds
                   7643
Seq. No.
                   9958_3.R1040
Contig ID
5'-most EST
                   jC-gmst02400031h06a1
                   BLASTN
Method
                                                                4,2
                   g410285
NCBI GI
BLAST score
                   90
E value
                   7.0e-43
Match length
                   186
% identity
                   87
NCBI Description Pisum sativum rho (ras-related) GTP-binding protein mRNA,
                   complete cds
Seq. No.
                   7644
Contig ID
                   9958 7.R1040
5'-most EST
                   eep7\overline{0}0869546.h1
Method
                   BLASTN
NCBI GI
                   q410285
BLAST score
                   42
E value
                   2.0e-14
Match length
                   62
% identity
NCBI Description Pisum sativum rho (ras-related) GTP-binding protein mRNA,
                   complete cds
Seq. No.
                   7645
Contig ID
                   9966 1.R1040
5'-most EST
                  asn701141112.h1
Seq. No.
                  7646
                   9966 2.R1040
Contig ID
5'-most EST
                  LIB3039-012-Q1-E1-G10
```

Seq. No. 7647 Contig ID 9967_1.R1040

5'-most EST LIB3039-012-Q1-E1-G12



```
BLASTX
Method
                  q1330343
NCBI GI
                  312
BLAST score
                  8.0e-29
E value
                  99
Match length
% identity
                  (U58755) C34D4.12 gene product [Caenorhabditis elegans]
NCBI Description
                  7648
Seq. No.
                  9979 1.R1040
Contig ID
5'-most EST
                  LIB3092-018-Q1-K1-D12
Method
                  BLASTX
                  g3212852
NCBI GI
                   387
BLAST score
                   2.0e-46
E value
                  176
Match length
% identity
                   53
                  (AC004005) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7649
Seq. No.
                   9988 1.R1040
Contig ID
5'-most EST
                   LIB3139-088-P1-N1-A7
                   BLASTX
Method
                   g4240116
NCBI GI
BLAST score
                   844
                   9.0e - 91
E value
Match length
                   189
% identity
                   (AB007799) NADH-cytochrome b5 reductase [Arabidopsis
NCBI Description
                   thaliana] >qi 4240118 dbj BAA74838_ (AB007800)
                   NADH-cytochrome b5 reductase [Arabidopsis thaliana]
Seq. No.
                   7650
                   9988 2.R1040
Contig ID
                   LIB3039-012-Q1-E1-C10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4336205
BLAST score
                   362
                   2.0e-34
E value
                   104
Match length
% identity
                   (AF077372) cytochrome b5 reductase [Zea mays]
NCBI Description
                   7651
Seq. No.
                   9989 2.R1040
Contig ID
                   zsg701128847.h1
5'-most EST
Method
                   BLASTX
                   g3894191
NCBI GI
BLAST score
                   423
                   2.0e-41
E value
                   158
Match length
                   54
% identity
                   (AC005662) unknown protein [Arabidopsis thaliana]
NCBI Description
```

Contig ID 9999_1.R1040

5'-most EST LIB3039-012-Q1-E1-D1

```
BLASTX
Method
                  g4006867
NCBI GI
                  741
BLAST score
                  3.0e-81
E value
                  197
Match length
                  81
% identity
                  (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
                  7653
Seq. No.
                  10010 1.R1040
Contig ID
                  ncj700979503.h2
5'-most EST
                  BLASTX
Method
                   g3860319
NCBI GI
                   1682
BLAST score
                   0.0e + 00
E value
                   442
Match length
                   77
% identity
                  (AJ012686) nucleolar protein [Cicer arietinum]
NCBI Description
                   7654
Seq. No.
                   10010 2.R1040
Contig ID
                   LIB3049-026-Q1-E1-F6
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4160346
                   337
BLAST score
                   2.0e-31
E value
                   142
Match length
                   50
% identity
                   (AL035216) nucleolar protein involved in pre-rRNA
NCBI Description
                   processing [Schizosaccharomyces pombe]
                   7655
Seq. No.
                   10010 3.R1040
Contig ID
                   LIB3139-026-P1-N1-C1
5'-most EST
                   BLASTN
Method
                   q3860318
NCBI GI
                   135
BLAST score
E value
                   1.0e-69
                   281
Match length
                   89
% identity
NCBI Description Cicer arietinum mRNA for nucleolar protein, partial
Seq. No.
                   7656
Contig ID
                   10018 1.R1040
                   trc700565153.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2462822
BLAST score
                   354
```

BLAST score 354 E value 4.0e-33 Match length 148

% identity 53

NCBI Description (AF000657) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7657

Contig ID 10027_1.R1040 5'-most EST jsh701063965.h1



```
7658
Seq. No.
                   10027 2.R1040
Contig ID
5'-most EST
                  ncj700979507.h2
Seq. No.
                   7659
                   10030 1.R1040
Contig ID
5'-most EST
                   kl1701213218.hl
                   BLASTX
Method
                   q2462758
NCBI GI
BLAST score
                   561
E value
                   3.0e-57
Match length
                   274
% identity
                   46
                   (AC002292) putative RNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7660
                   10040 1.R1040
Contig ID
                   fua701038228.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3688172
BLAST score
                   176
                   2.0e-12
E value
Match length
                   64
% identity
                   62
                  (AL031804) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7661
                   10040 2.R1040
Contig ID
                   LIB3109-052-Q1-K1-B3
5'-most EST
                   BLASTX
Method
                   g3688172
NCBI GI
                   176
BLAST score
                   1.0e-12
E value
Match length
                   64
% identity
                   64
                  (AL031804) putative protein [Arabidopsis thaliana]
NCBI Description
                   7662
Seq. No.
Contig ID
                   10040 3.R1040
5'-most EST
                   trc700566501.h1
Seq. No.
                   7663
                   10048 1.R1040
Contig ID
5'-most EST
                   txt700737261.hl
                   7664
Seq. No.
                   10062 1.R1040
Contig ID
5'-most EST
                   LIB3040-050-Q1-E1-D1
                   BLASTX
Method
```

Method BLASTX
NCBI GI g2842662
BLAST score 176
E value 1.0e-12
Match length 49
% identity 59

NCBI Description DNA-DIRECTED RNA POLYMERASES I, II, AND III 7.0 KD

POLYPEPTIDE (ABC10-ALPHA) (RPB7.0) (METALLOTHIONEIN-I GENE



TRANSCRIPTION ACTIVATOR)

```
Seq. No.
                  7665
                  10068 1.R1040
Contig ID
5'-most EST
                  LIB3109-038-Q1-K1-A10
Seq. No.
                  7666
                  10068 4.R1040
Contig ID
5'-most EST
                  zhf700964470.hl
                  BLASTX
Method
NCBI GI
                  g3876716
BLAST score
                  206
E value
                  2.0e-16
Match length
                  66
% identity
                  48
NCBI Description (Z46242) similar to Zinc finger, C3HC4 type (RING finger)
                  [Caenorhabditis elegans]
                  7667
Seq. No.
Contig ID
                  10071 1.R1040
5'-most EST
                  uC-gmropic115d03b1
Seq. No.
                  7668
                  10071 2.R1040
Contig ID
5'-most EST
                  leu701157872.hl
Method
                  BLASTN
NCBI GI
                  g532289
BLAST score
                  49
                  2.0e-18
E value
Match length
                  61
                  95
% identity
NCBI Description Soybean mRNA for late nodulin, complete cds
Seq. No.
                  7669
Contig ID
                  10071_3.R1040
5'-most EST
                  uC-gmropic014c09b1
Method
                  BLASTN
NCBI GI
                  g532289
BLAST score
                  76
E value
                  2.0e-34
Match length
                  84
                  98
% identity
NCBI Description Soybean mRNA for late nodulin, complete cds
Seq. No.
                  7670
                  10073 1.R1040
Contig ID
5'-most EST
                  LIB3039-011-Q1-E1-B2
Method
                  BLASTX
NCBI GI
                  g2262104
BLAST score
                  439
```

E value 1.0e-43 Match length 121

% identity 69

NCBI Description (AC002343) unknown protein [Arabidopsis thaliana]

Seq. No.

7671

10076 1.R1040 Contig ID



```
LIB3039-011-Q1-E1-B5
5'-most EST
                   7672
Seq. No.
                   10078 1.R1040
Contig ID
5'-most EST
                   gsv701053747.hl
Method
                   BLASTX
                   q3021409
NCBI GI
                   212
BLAST score
                   8.0e-17
E value
                   70
Match length
% identity
NCBI Description (Y12781) transducin (beta) like 1 protein [Homo sapiens]
Seq. No.
                   10098 1.R1040
Contig ID
5'-most EST
                   LIB3051-005-Q1-E1-C6
Method
                   BLASTX
NCBI GI
                   q3193298
                   214
BLAST score
                   2.0e-35
E value
                   137
Match length
                   55
% identity
NCBI Description (AF069298) T14P8.17 gene product [Arabidopsis thaliana]
                   7674
Seq. No.
                   10098 2.R1040
Contig ID
5'-most EST
                   zhf700955551.h1
                   7675
Seq. No.
                   10100 1.R1040
Contig ID
                   LIB3039-010-Q1-E1-G5
5'-most EST
Method
                   BLASTN
                   q1732468
NCBI GI
BLAST score
                   64
                   2.0e-27
E value
                   136
Match length
% identity
                   87
                   Soybean mRNA for Mg chelatase subunit (46 kD), complete cds
NCBI Description
                   7676
Seq. No.
                   10105 1.R1040
Contig ID
                   LIB30\overline{3}9-010-Q1-E1-E11
5'-most EST
                   BLASTX
Method
                   g3608154
NCBI GI
                   288
BLAST score
                   7.0e-26
E value
Match length
                   137
                   43
% identity
                   (AC005314) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7677
 Seq. No.
                   10107 1.R1040
 Contig ID
                   LIB3053-002-Q1-B1-C4
 5'-most EST
                   BLASTX
Method
                   q4539422
NCBI GI
BLAST score
                    1186
```

1.0e-130

E value

```
306
Match length
                  74
% identity
NCBI Description (AL049171) putative protein [Arabidopsis thaliana]
Seq. No.
                  7678
Contig ID
                  10120 1.R1040
                  LIB3039-010-Q1-E1-D1
5'-most EST
                  BLASTX
Method
                  g498040
NCBI GI
BLAST score
                  315
                  4.0e-29
E value
Match length
                  94
                  66
% identity
NCBI Description (L33793) ORF [Senecio odorus]
                  7679
Seq. No.
Contig ID
                  10127 1.R1040
5'-most EST
                  q4302191
Method
                  BLASTX
                  q730463
NCBI GI
                  310
BLAST score
                  3.0e-28
E value
                  105
Match length
                  55
% identity
                  60S RIBOSOMAL PROTEIN L37B (YL37) (RP47)
NCBI Description
                  >gi_630323_pir__S44069 ribosomal protein L35a.e.c15 - yeast
                   (Saccharomyces cerevisiae) >gi_484241 (L23923) ribosomal
                  protein L37 [Saccharomyces cerevisiae]
                   >gi_1420537_emb_CAA99454_ (Z75142) ORF YOR234c
                   [Saccharomyces cerevisiae]
                   7680
Seq. No.
                   10130 1.R1040
Contig ID
                  hrw701063208.h1
5'-most EST
```

Contig ID 5'-most EST 10130 2.R1040 sat701012105.h1

Seq. No. 7682

Contig ID 10144_1.R1040

5'-most EST LIB3109-052-Q1-K1-B5

Method BLASTX
NCBI GI g135053
BLAST score 705
E value 3.0e-74
Match length 302
% identity 47

NCBI Description SULFITE OXIDASE >gi_86440_pir__A34180 sulfite oxidase (EC

1.8.3.1), hepatic - chicken

Seq. No. 7683

Contig ID 10146_1.R1040

5'-most EST LIB3039-010-Q1-E1-A2

Seq. No. 7684

Contig ID 10158_1.R1040

1410

Contig ID 5'-most EST

Method



```
LIB3039-009-Q1-E1-G8
5'-most EST
                  BLASTX
Method
                  g3335359
NCBI GI
                  337
BLAST score
                   1.0e-31
E value
Match length
                   92
% identity
                  (AC003028) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7685
Seq. No.
Contig ID
                   10159 1.R1040
                   uC-gmropic087d11b1
5'-most EST
                   BLASTX
Method
                   q2190187
NCBI GI
                   194
BLAST score
                   2.0e-14
E value
Match length
                   201
                   33
% identity
                  (D64087) nuclear matrix constituent protein 1 (NMCP1)
NCBI Description
                   [Daucus carota]
                   7686
Seq. No.
                   10166 1.R1040
Contig ID
                   LIB3039-009-Q1-E1-F10
5'-most EST
                   7687
Seq. No.
Contig ID
                   10179 1.R1040
                   LIB3138-028-Q1-N1-G10
5'-most EST
                   BLASTX
Method
                   q4455326
NCBI GI
BLAST score
                   621
                   2.0e-64
E value
                   196
Match length
                   60
% identity
                  (AL035525) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7688
                   10180 1.R1040
Contig ID
                   sat701003879.h1
5'-most EST
                   BLASTX
Method
                   q421843
NCBI GI
BLAST score
                   1396
                   1.0e-155
E value
                   315
Match length
                   84
 % identity
                   protein kinase (EC 2.7.1.37) 5 - Arabidopsis thaliana
NCBI Description
                   >gi_217861_dbj_BAA01715_ (D10909) serine/threonine protein
                   kinase [Arabidopsis thaliana]
                   7689
 Seq. No.
                   10183 1.R1040
 Contig ID
                   fC-gmse700675307a1
 5'-most EST
                    7690
 Seq. No.
                    10183 2.R1040
```

1411

leu701148361.h1

BLASTX



```
g1280434
NCBI GI
                  178
BLAST score
                  5.0e-13
E value
                  117
Match length
                  38
% identity
                  (U42014) hemomucin [Drosophila melanogaster]
NCBI Description
                  7691
Seq. No.
                  10188 1.R1040
Contig ID
                  LIB3053-008-Q1-N1-A10
5'-most EST
                  BLASTX
Method
                   g3894168
NCBI GI
                   503
BLAST score
                   9.0e-51
E value
                   190
Match length
                   52
% identity
                   (AC005312) similar to phloem-specific lectin [Arabidopsis
NCBI Description
                   thaliana]
                   7692
Seq. No.
                   10188 2.R1040
Contig ID
                   LIB3049-055-Q1-E1-B2
5'-most EST
                   BLASTX
Method
                   g3894158
NCBI GI
BLAST score
                   218
                   1.0e-22
E value
                   131
Match length
                   44
% identity
                   (AC005312) similar to phloem-specific lectin [Arabidopsis
NCBI Description
                   thaliana]
                   7693
Seq. No.
Contig ID
                   10206 1.R1040
                   LIB3039-008-Q1-E1-G6
5'-most EST
                   7694
Seq. No.
                   10207 1.R1040
Contig ID
5'-most EST
                   g4303898
Method
                   BLASTN
                   g170643
NCBI GI
BLAST score
                   301
                   1.0e-168
E value
Match length
                   440
                   94
 % identity
                   Vigna aconitifolia
NCBI Description
                   5'-phosphoribosyl-4-(N-succinocarboxamide)-5-
                   aminoimidazole synthetase mRNA
                   7695
 Seq. No.
                   10207 2.R1040
 Contig ID
                   xpa700798252.h1
 5'-most EST
                   BLASTN
 Method
                   g170643
 NCBI GI
                   77
 BLAST score
                   4.0e-35
 E value
```

225

93

Match length % identity



```
Vigna aconitifolia
NCBI Description
                  5'-phosphoribosyl-4-(N-succinocarboxamide)-5-
                  aminoimidazole synthetase mRNA
Seq. No.
                  7696
                  10213 1.R1040
Contig ID
5'-most EST
                  awf700838202.h1
                  BLASTX
Method
                  g2213614
NCBI GI
BLAST score
                  268
E value
                  2.0e-23
Match length
                  116
% identity
                  (AC000103) F21J9.8 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  7697
                   10221 1.R1040
Contig ID
                  LIB3167-010-P1-K1-G8
5'-most EST
                   7698
Seq. No.
                   10221 2.R1040
Contig ID
5'-most EST
                  wvk700681917.hl
                   7699
Seq. No.
Contig ID
                   10224 1.R1040
5'-most EST
                   leu701151970.h1
Method
                   BLASTX
NCBI GI
                   g2462834
BLAST score
                   206
                   3.0e-16
E value
                   121
Match length
% identity
                   40
                  (AF000657) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   7700
Seq. No.
                   10225 1.R1040
Contig ID
                   LIB3040-008-Q1-E1-A3
5'-most EST
                   7701
Seq. No.
                   10228 1.R1040
Contig ID
5'-most EST
                   LIB3039-008-Q1-E1-E8
                   BLASTX
Method
                   g541950
NCBI GI
BLAST score
                   160
                   7.0e-11
E value
Match length
                   80
                   46
% identity
NCBI Description SPCP1 protein - soybean >gi 310576 (L12257) nodulin-26
                   [Glycine max]
```

10235 1.R1040 Contig ID

5'-most EST jC-qmle01810065b01a1

BLASTX Method NCBI GI q4539420 BLAST score 224 E value 3.0e-18



```
Match length
                  80
                  61
% identity
                  (ALO49171) putative protein (fragment) [Arabidopsis
NCBI Description
                  thaliana]
                  7703
Seq. No.
                  10249 2.R1040
Contig ID
                  uxk700672272.h1
5'-most EST
                  7704
Seq. No.
Contig ID
                  10250 1.R1040
5'-most EST
                  LIB3107-005-Q1-K1-F9
Method
                  BLASTX
NCBI GI
                  g3025188
BLAST score
                  366
                  1.0e-34
E value
Match length
                  107
                  61
% identity
NCBI Description HYPOTHETICAL 11.9 KD PROTEIN SLR1846
                  >gi 1652154 dbj BAA17078 (D90903) hypothetical protein
                  [Synechocystis sp.]
                  7705
Seq. No.
Contig ID
                  10261 1.R1040
5'-most EST
                  ncj700978977.h1
Method
                  BLASTX
NCBI GI
                  g1362615
BLAST score
                  810
E value
                  2.0e-86
Match length
                  404
% identity
                  44
NCBI Description iswi protein - fruit fly (Drosophila melanogaster)
                  7706
Seq. No.
```

>gi 439197 (L27127) ISWI protein [Drosophila melanogaster]

Contig ID 10264 1.R1040

LIB3039-008-Q1-E1-A10 5'-most EST

BLASTX Method NCBI GI q2760326 BLAST score 145 E value 3.0e-09 Match length 88 % identity

NCBI Description (AC002130) F1N21.11 [Arabidopsis thaliana]

Seq. No. 7707

10269 1.R1040 Contig ID 5'-most EST gsv701044623.h1

Method BLASTX NCBI GI g2911799 BLAST score 724 E value 1.0e-76 Match length 222 % identity 61

NCBI Description (AF008184) 4-coumarate: CoA ligase 1 [Populus balsamifera

subsp. trichocarpa X Populus deltoides]

NCBI Description

(DP-3)



7708 Seq. No. 10276 1.R1040 Contig ID LIB3040-060-Q1-E1-F6 5'-most EST 7709 Seq. No. Contig ID 10284 1.R1040 LIB3049-001-Q1-E1-A1 5'-most EST Method BLASTX q4539417 NCBI GI 357 BLAST score 2.0e-33 E value 106 Match length % identity NCBI Description (AL049171) putative protein [Arabidopsis thaliana] Seq. No. 10284 2.R1040 Contig ID 5'-most EST LIB3106-041-Q1-K1-A7 BLASTX Method a4539417 NCBI GI 511 BLAST score E value 1.0e-51 121 Match length 76 % identity NCBI Description (AL049171) putative protein [Arabidopsis thaliana] Seq. No. 7711 Contig ID 10289 1.R1040 5'-most EST ssr700554183.hl Seq. No. 7712 10289 2.R1040 Contig ID uC-gmropic029g05b1 5'-most EST 7713 Seq. No. 10298 1.R1040 Contig ID $LIB30\overline{3}9-007-Q1-E1-A9$ 5'-most EST BLASTX Method g913227 NCBI GI 491 BLAST score 3.0e-49 E value 200 Match length 53 % identity transcription factor XlDP-1 [Xenopus laevis, embryos, NCBI Description Peptide, 409 aa] 7714 Seq. No. 10298 2.R1040 Contig ID 5'-most EST $vzy70\overline{0}755726.h1$ BLASTX Method NCBI GI q3122929 170 BLAST score 8.0e-12 E value 63 Match length % identity

1415

TRANSCRIPTION FACTOR DP-2 (E2F DIMERIZATION PARTNER 2)



```
7715
Seq. No.
                   10298 3.R1040
Contig ID
                  pcp700994367.hl
5'-most EST
                   7716
Seq. No.
                   10312 1.R1040
Contig ID
                   LIB30\overline{5}1-080-Q1-K1-G2
5'-most EST
                   BLASTN
Method
                   g2618602
NCBI GI
                   41
BLAST score
                   2.0e-13
E value
                   104
Match length
                   85
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MSJ1, complete sequence [Arabidopsis thaliana]
                   7717
Seq. No.
                   10318 1.R1040
Contig ID
                   LIB3039-007-Q1-E1-A2
5'-most EST
                   BLASTX
Method
                   g4455202
NCBI GI
                   199
BLAST score
                   2.0e-15
E value
                   55
Match length
                   62
% identity
                   (AL035440) putative APG protein [Arabidopsis thaliana]
NCBI Description
                   7718
Seq. No.
                   10328 1.R1040
Contig ID
                   smc700749713.hl
5'-most EST
                   7719
Seq. No.
                   10341 1.R1040
Contig ID
                   trc700564224.hl
5'-most EST
                   BLASTX
Method
                   g3445208
NCBI GI
BLAST score
                   601
                   2.0e-62
E value
                   146
Match length
                   78
% identity
                   (AC004786) putative amino-acid acetyltransferase
NCBI Description
                   [Arabidopsis thaliana]
                   7720
Seq. No.
                   10342 1.R1040
Contig ID
                   fC-gmle700787001a2
5'-most EST
                   7721
Seq. No.
                   10347 1.R1040
Contig ID
                   LIB3039-006-Q1-E1-E7
5'-most EST
                   BLASTX
Method
                   g4335763
NCBI GI
```

1416

329

91

70

9.0e-31

BLAST score

Match length

% identity

E value



NCBI Description (AC006284) unknown protein [Arabidopsis thaliana]

Seq. No. 7722

Contig ID 10349 1.R1040

5'-most EST LIB3039-006-Q1-E1-B11

Seq. No. 7723 Contig ID 10358_1.R1040 5'-most EST LIB3106-069-P1-K1-A1

Method BLASTX
NCBI GI g3980254
BLAST score 600
E value 5.0e-62
Match length 183
% identity 61

NCBI Description (AJ006053) peroxisomal membrane protein [Arabidopsis

thaliana]

Seq. No. 7724

Contig ID 10358_5.R1040 5'-most EST g5753066

Seq. No. 7725

Contig ID 10360_1.R1040 5'-most EST epx701107133.h1

Method BLASTX
NCBI GI g4538929
BLAST score 356
E value 9.0e-34
Match length 90
% identity 74

NCBI Description (AL049483) putative nucleic acid binding protein

[Arabidopsis thaliana]

Seq. No. 7726

Contig ID 10364_1.R1040

5'-most EST LIB3039-005-Q1-E1-H11

Method BLASTX
NCBI GI g3702323
BLAST score 185
E value 1.0e-13
Match length 83
% identity 48

NCBI Description (AC005397) unknown protein [Arabidopsis thaliana]

Seq. No. 7727

Contig ID 10375_1.R1040 5'-most EST leu701153565.h1

Seq. No. 7728

Contig ID 10389_1.R1040

5'-most EST LIB3040-008-Q1-E1-H1

Method BLASTX
NCBI GI g1653767
BLAST score 1176
E value 1.0e-129
Match length 400

1417



```
% identity
NCBI Description (D90916) oligopeptidase A [Synechocystis sp.]
                  7729
Seq. No.
                  10391 1.R1040
Contig ID
5'-most EST
                  pxt700944410.h1
Method
                  BLASTN
                  g166409
NCBI GI
                  235
BLAST score
                  1.0e-129
E value
                  423
Match length
% identity
                  89
                  Alfalfa nucleic acid binding protein (alfin-1) mRNA,
NCBI Description
                  partial cds
                  7730
Seq. No.
                  10391 2.R1040
Contig ID
                  sat701005773.h1
5'-most EST
                  BLASTN
Method
NCBI GI
                  g166409
                  135
BLAST score
                  5.0e-70
E value
                  203
Match length
                  92
% identity
                  Alfalfa nucleic acid binding protein (alfin-1) mRNA,
NCBI Description
                  partial cds
                  7731
Seq. No.
                  10391 3.R1040
Contig ID
                   zsq701130201.h1
5'-most EST
                  BLASTN
Method
                   q166409
NCBI GI
BLAST score
                   65
                   4.0e-28
E value
                   153
Match length
% identity
                   86
                  Alfalfa nucleic acid binding protein (alfin-1) mRNA,
NCBI Description
                  partial cds
                   7732
Seq. No.
                   10394 1.R1040
Contig ID
5'-most EST
                   LIB3039-005-Q1-E1-D9
Method
                   BLASTX
NCBI GI
                   q728868
BLAST score
                   239
                   4.0e-20
E value
Match length
                   79
% identity
                   58
                   ANTER-SPECIFIC PROLINE-RICH PROTEIN APG (PROTEIN CEX)
NCBI Description
                   >gi_99824_pir__S16748 proline-rich protein - rape
                   (fragment) >gi_22597_emb_CAA42924_ (X60376) proline-rich
                   protein [Brassica napus]
```

Contig ID 10394_2.R1040 5'-most EST zhf700954572.h1

Method BLASTX



```
q728868
NCBI GI
                  140
BLAST score
                  1.0e-08
E value
                  44
Match length
% identity
                  ANTER-SPECIFIC PROLINE-RICH PROTEIN APG (PROTEIN CEX)
NCBI Description
                  >gi 99824 pir__S16748 proline-rich protein - rape
                  (fragment) >gi_22597_emb_CAA42924_ (X60376) proline-rich
                  protein [Brassica napus]
                  7734
Seq. No.
                  10395 1.R1040
Contig ID
                  LIB3049-016-Q1-E1-H10
5'-most EST
                  BLASTX
Method
                   q2618702
NCBI GI
                   461
BLAST score
                   5.0e-46
E value
                   143
Match length
% identity
                   (AC002510) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7735
Seq. No.
                   10396 1.R1040
Contig ID
                   LIB3051-062-Q1-K1-E12
5'-most EST
                   BLASTX
Method
                   q4090257
NCBI GI
BLAST score
                   467
                   8.0e-47
E value
                   91
Match length
% identity
                   (AJ131732) ribosomal protein L37A [Pseudotsuga menziesii]
NCBI Description
Seq. No.
                   7736
                   10396 2.R1040
Contig ID
                   LIB3039-005-Q1-E1-E11
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4090257
BLAST score
                   452
                   7.0e-45
E value
                   91
Match length
% identity
                   96
                   (AJ131732) ribosomal protein L37A [Pseudotsuga menziesii]
NCBI Description
                   7737
Seq. No.
                   10410 1.R1040
Contig ID
                   g5606468
 5'-most EST
                   BLASTX
Method
                   g3121825
NCBI GI
                   994
BLAST score
                   1.0e-108
 E value
                   260
Match length
                   77
 % identity
                   2-CYS PEROXIREDOXIN BAS1 PRECURSOR (THIOL-SPECIFIC
 NCBI Description
                   ANTIOXIDANT PROTEIN) >gi_1498247_emb_CAA63910_ (X94219)
```

bas1 protein [Spinacia oleracea]



```
Contiq ID
                  10410 2.R1040
5'-most EST
                  LIB3039-005-Q1-E1-D11
                  BLASTX
Method
NCBI GI
                  g3121825
BLAST score
                  619
                  2.0e-64
E value
Match length
                  182
                  70
% identity
NCBI Description 2-CYS PEROXIREDOXIN BAS1 PRECURSOR (THIOL-SPECIFIC
                  ANTIOXIDANT PROTEIN) >gi_1498247_emb_CAA63910_ (X94219)
                  bas1 protein [Spinacia oleracea]
                  7739
Seq. No.
                  10418 1.R1040
Contig ID
5'-most EST
                  LIB3039-005-Q1-E1-B10
Method
                  BLASTN
NCBI GI
                  g1495767
BLAST score
                  238
E value
                  1.0e-131
Match length
                  611
% identity
                  84
NCBI Description P.sativum mRNA for 110 kD chloroplast inner envelope
                  protein IEP110
                  7740
Seq. No.
                  10425 1.R1040
Contig ID
                  leu701157388.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2897942
BLAST score
                  317
E value
                  7.0e-29
Match length
                  165
% identity
                  42
                  (AF003938) thioredoxin-like protein [Homo sapiens]
NCBI Description
                  >gi_2961254 (AF051896) thioredoxin homolog [Homo sapiens]
                  >gi_2970689 (AF052659) thioredoxin-related protein [Homo
                  sapiens]
                  7741
Seq. No.
                  10428 1.R1040
Contig ID
5'-most EST
                  LIB3049-026-Q1-E1-G3
Method
                  BLASTX
NCBI GI
                  g4538959
BLAST score
                  460
                  7.0e-46
E value
                  105
Match length
% identity
NCBI Description (AL049488) putative protein [Arabidopsis thaliana]
Seq. No.
                  7742
Contig ID
                  10430 1.R1040
```

5'-most EST LIB3039-004-Q1-E1-H1

7743 Seq. No.

10432 1.R1040 Contig ID 5'-most EST rlr700902245.hl

Method BLASTX



```
g2618688
NCBI GI
                   681
BLAST score
                   1.0e-71
E value
                   157
Match length
                   82
% identity
                  (AC002510) putative esterase D [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10435 1.R1040
Contig ID
                   LIB3051-116-Q1-K1-A2
5'-most EST
                   7745
Seq. No.
                   10439 1.R1040
Contig ID
                   leu701144983.h1
5'-most EST
                   BLASTX
Method
                   g4191782
NCBI GI
                   2171
BLAST score
                   0.0e + 00
E value
                   503
Match length
                   82
% identity
NCBI Description (AC005917) WD-40 repeat protein [Arabidopsis thaliana]
                   7746
Seq. No.
                   10449 1.R1040
Contig ID
                   LIB3039-004-Q1-E1-D9
5'-most EST
                   BLASTX
Method
                   g4263517
NCBI GI
BLAST score
                   199
                   2.0e-31
E value
                   112
Match length
% identity
                   64
                   (AC004044) similar to PHZF, catalyzing the hydroxylation of
NCBI Description
                   phenazine-1-carboxylic acid to
                   2-hydroxy-phenazine-1-carboxylic acid [Arabidopsis
                   thaliana]
                   7747
Seq. No.
Contig ID
                   10453 1.R1040
                   LIB3040-043-Q1-E1-H7
5'-most EST
                   BLASTX
Method
                    g4406774
NCBI GI
BLAST score
                    920
E value
                    2.0e-99
Match length
                    264
 % identity
                    64
                    (AC006836) putative nonsense-mediated mRNA decay protein,
NCBI Description
                    5' partial [Arabidopsis thaliana]
                    7748
 Seq. No.
                    10458 1.R1040
 Contig ID
                    seb70\overline{0}652912.h1
 5'-most EST
                    BLASTX
 Method
                    g3138972
 NCBI GI
                    258
 BLAST score
                    3.0e-22
 E value
                    67
 Match length
```

73

% identity



```
(AF038505) dihydrolipoylacyltransferase subunit of the
NCBI Description
                  branched-chain alpha-keto acid dehydrogenase complex
                  [Arabidopsis thaliana]
                  7749
Seq. No.
                  10462 1.R1040
Contig ID
                  LIB3092-010-Q1-K1-C3
5'-most EST
                  7750
Seq. No.
                  10470 1.R1040
Contig ID
                  LIB3050-022-Q1-K1-A3
5'-most EST
                  BLASTX
Method
                  g2131767
NCBI GI
                   165
BLAST score
                   3.0e-11
E value
                   96
Match length
% identity
                   hypothetical protein YLR074c - yeast (Saccharomyces
NCBI Description
                   cerevisiae) >gi_1360432_emb_CAA97631_ (Z73246) ORF YLR074c
                   [Saccharomyces cerevisiae]
                   7751
Seq. No.
                   10476 1.R1040
Contig ID
                   LIB3109-020-Q1-K1-B3
5'-most EST
                   BLASTX
Method
                   g2832680
NCBI GI
                   476
BLAST score
                   8.0e-48
E value
                   107
Match length
                   40
% identity
                   (AL021712) putative protein [Arabidopsis thaliana]
NCBI Description
                   7752
Seq. No.
                   10490 1.R1040
Contig ID
5'-most EST
                   pcp700993962.hl
                   7753
Seq. No.
                   10491 1.R1040
Contig ID
                   LIB3039-003-Q1-E1-F7
5'-most EST
Method
                   BLASTX
                   q4415906
NCBI GI
BLAST score
                   302
E value
                   2.0e-27
Match length
                   96
 % identity
                   (AC006282) putative DNA binding protein with homeobox
NCBI Description
                   domain [Arabidopsis thaliana]
 Seq. No.
                   7754
                   10492 1.R1040
 Contig ID
                   LIB3039-003-Q1-E1-F8
 5'-most EST
 Method
                   BLASTX
 NCBI GI
                   g1709543
```

229 BLAST score 1.0e-18 E value

95 Match length 48 % identity



```
RED CELL ACID PHOSPHATASE 1, ISOZYME F (ACP1) (LOW
NCBI Description
                  MOLECULAR WEIGHT PHOSPHOTYROSINE PROTEIN PHOSPHATASE)
                  (ADIPOCYTE ACID PHOSPHATASE, ISOZYME ALPHA)
                  >gi_284353_pir__A38148 protein-tyrosine-phosphatase (EC
                  3.1.3.48), low molecular weight, splice form f - human
                  >gi 179636 (M83653) cytoplasmic phosphotyrosyl protein
                  phosphatase [Homo sapiens] >gi_1147812 (U25849) red
                  cell-type low molecular weight acid phosphatase [Homo
                  sapiens]
Seq. No.
                  7755
Contig ID
                  10493 1.R1040
5'-most EST
                  dpv701100495.h1
                  BLASTX
Method
                  g1731990
NCBI GI
BLAST score
                  283
E value
                  1.0e-47
Match length
                  247
% identity
                  43
                  (Y09602) serine carboxypeptidase II, CP-MII [Hordeum
NCBI Description
                  vulgare]
                  7756
Seq. No.
Contig ID
                  10495 1.R1040
                  LIB3039-003-Q1-E1-G12
5'-most EST
                  BLASTX
Method
NCBI GI
                  g4049341
BLAST score
                   402
E value
                   5.0e-40
Match length
                  176
                   49
% identity
                  (AL034567) putative protein [Arabidopsis thaliana]
NCBI Description
                   7757
Seq. No.
                   10496 1.R1040
Contig ID
                   jC-gmf102220148c06a1
5'-most EST
Method
                   BLASTX
                   q3342556
NCBI GI
BLAST score
                   1960
                   0.0e + 00
E value
                   480
Match length
                   80
% identity
                  (AF077528) importin alpha [Arabidopsis thaliana]
NCBI Description
                   7758
Seq. No.
                   10496_2.R1040
Contig ID
                   jC-gmro02910067a11a1
5'-most EST
                   BLASTX
Method
                   q3273243
NCBI GI
                   552
```

Method BLASTX
NCBI GI g3273243
BLAST score 552
E value 1.0e-56
Match length 167
% identity 39

NCBI Description (AB004660) NLS receptor [Oryza sativa]

>gi 3273245 dbj_BAA31166_ (AB004814) NLS receptor [Oryza

satīva]



```
7759
Seq. No.
                  10496_3.R1040
Contig ID
                  uC-gmropic073g11b1
5'-most EST
                  BLASTX
Method
                  g2950210
NCBI GI
                   375
BLAST score
                   6.0e-36
E value
                   102
Match length
% identity
                   81
                  (Y14615) Importin alpha-like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10496 7.R1040
Contig ID
                   gsv701054894.hl
5'-most EST
                   {\tt BLASTX}
Method
                   g2950210
NCBI GI
                   237
BLAST score
                   6.0e-20
E value
                   68
Match length
                   75
% identity
                   (Y14615) Importin alpha-like protein [Arabidopsis thaliana]
NCBI Description
                   7761
Seq. No.
                   10500 1.R1040
Contig ID
                   crh700856228.hl
5'-most EST
                   7762
Seq. No.
                   10502 1.R1040
Contig ID
                   LIB3039-001-Q1-E1-H9
5'-most EST
                   BLASTX
Method
                   g3341443
NCBI GI
                   307
BLAST score
                   4.0e-28
E value
                   99
Match length
% identity
                   (AJ223074) acid phosphatase [Glycine max]
NCBI Description
                   7763
Seq. No.
                   10503 1.R1040
Contig ID
                   LIB3107-009-Q1-K1-F8
5'-most EST
                   BLASTX
Method
NCBI GI
                    q4454466
BLAST score
                    203
E value
                    1.0e-15
                    73
Match length
                    56
 % identity
                   (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                    7764
 Seq. No.
                    10504 1.R1040
 Contig ID
                    LIB3039-003-Q1-E1-E12
 5'-most EST
```

10517 1.R1040 Contig ID

LIB3109-056-Q1-K1-E5 5'-most EST

BLASTX Method g3107903 NCBI GI

```
279
BLAST score
                  1.0e-24
E value
                  104
Match length
                  51
% identity
NCBI Description (D83719) polycomb-like protein [Daucus carota]
                  7766
Seq. No.
                  10520 1.R1040
Contig ID
                  rlr700897049.hl
5'-most EST
                  BLASTX
Method
                  q470373
NCBI GI
BLAST score
                  306
                  2.0e-27
E value
                  165
Match length
                  39
% identity
                  (U00047) ZK418.5 gene product [Caenorhabditis elegans]
NCBI Description
                  7767
Seq. No.
                  10520 2.R1040
Contig ID
                  LIB3170-079-Q1-K1-F2
5'-most EST
                  BLASTX
Method
                   g470373
NCBI GI
                   201
BLAST score
                   2.0e-15
E value
                   117
Match length
                   38
% identity
NCBI Description (U00047) ZK418.5 gene product [Caenorhabditis elegans]
                   7768
Seq. No.
                   10530 1.R1040
Contig ID
                   LIB3049-007-Q1-E1-F1
5'-most EST
                   7769
Seq. No.
                   10540 1.R1040
Contig ID
                   LIB3072-017-Q1-E1-A6
5'-most EST
                   BLASTX
Method
                   g4039152
NCBI GI
BLAST score
                   175
                   2.0e-12
E value
                   36
Match length
                   81
% identity
                   (AF104221) low temperature and salt responsive protein
NCBI Description
                   LTI6B [Arabidopsis thaliana] >gi_4325219_gb_AAD17303
                   (AF122006) hydrophobic protein [Arabidopsis thaliana]
                   7770
Seq. No.
                   10541 1.R1040
Contig ID
                   LIB3055-002-Q1-B1-G7
5'-most EST
                   BLASTN
Method
                   g3885510
NCBI GI
                   157
BLAST score
                   1.0e-82
E value
                   332
Match length
                   89
 % identity
NCBI Description Medicago sativa clone MS28 unknown mRNA
```



Contig ID 10541_3.R1040

5'-most EST LIB3039-003-Q1-E1-A12

Method BLASTN
NCBI GI g3885510
BLAST score 94
E value 4.0e-45
Match length 228
% identity 87

NCBI Description Medicago sativa clone MS28 unknown mRNA

Seq. No. 7772

Contig ID 10546_1.R1040

5'-most EST LIB3039-002-Q1-E1-H12

Method BLASTX
NCBI GI g3548810
BLAST score 277
E value 2.0e-36
Match length 279
% identity 37

NCBI Description (AC005313) putative chloroplast nucleoid DNA binding

protein [Arabidopsis thaliana]

Seq. No. 7773

Contig ID 10560_1.R1040 5'-most EST cf1700863674.h1

Method BLASTX
NCBI GI g3776560
BLAST score 484
E value 9.0e-49
Match length 145
% identity 69

NCBI Description (AC005388) Similar to gb_U51990 hPrp18 (splicing factor)

gene from Homo sapiens. [Arabidopsis thaliana]

Seq. No. 7774

Contig ID 10561_1.R1040

5'-most EST LIB3106-070-P1-K1-A1

Method BLASTX
NCBI GI g548702
BLAST score 391
E value 9.0e-38
Match length 111
% identity 58

NCBI Description DNA-DIRECTED RNA POLYMERASE II 14.5 KD POLYPEPTIDE (RPB9)

(RPB14.5) >gi_543001_pir__S41621 DNA-directed RNA polymerase (EC 2.7.7.6) II 14.5K chain - human

>gi_397150_emb_CAA80649_ (Z23102) RNA Polymerase II subunit

14.5 kD [Homo sapiens] >gi 1905901 (AD001527) HUMAN DNA-DIRECTED RNA POLYMERASE II 14.5 KD SUBUNIT [Homo

sapiens]

Seq. No. 7775

Contig ID 10571_1.R1040 5'-most EST sat701006018.h2

Method BLASTX
NCBI GI g3582339
BLAST score 580



```
1.0e-59
E value
Match length
                  141
                  74
% identity
NCBI Description (AC005496) unknown protein [Arabidopsis thaliana]
                  7776
Seq. No.
                  10585 1.R1040
Contig ID
                   jC-gm\overline{l}e01810089e09a1
5'-most EST
                  BLASTX
Method
                  g3785983
NCBI GI
BLAST score
                   312
                   1.0e-28
E value
                   108
Match length
                   56
% identity
NCBI Description (AC005560) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   7777
                   10589 1.R1040
Contig ID
5'-most EST
                   LIB3039-002-Q1-E1-D5
                   BLASTX
Method
                   g4539345
NCBI GI
BLAST score
                   355
                   7.0e-34
E value
Match length
                   94
                   71
% identity
                  (AL035539) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   7778
Seq. No.
                   10591 1.R1040
Contig ID
                   LIB3039-002-Q1-E1-A2
5'-most EST
                   7779
Seq. No.
                   10591 2.R1040
Contig ID
                   uC-gmropic045a04b1
5'-most EST
                   7780
Seq. No.
                   10596 1.R1040
Contig ID
                   LIB3092-055-Q1-K1-E4
5'-most EST
                   BLASTX
Method
                   g2894612
NCBI GI
                   2489
BLAST score
                   0.0e + 00
E value
                   613
Match length
                   74
% identity
                   (AL021889) putative protein [Arabidopsis thaliana]
NCBI Description
                   7781
Seq. No.
                   10596 2.R1040
Contig ID
5'-most EST
                   ncj700986073.hl
                   BLASTX
Method
NCBI GI
                   q4262174
                   513
BLAST score
                   4.0e-52
E value
                   129
Match length
% identity
```

NCBI Description (AC005508) 9058 [Arabidopsis thaliana]



7782 Seq. No. 10596 3.R1040 Contig ID LIB3092-055-Q1-K1-E6 5'-most EST BLASTX Method q4262174 NCBI GI BLAST score 197 4.0e-15 E value 69 Match length % identity NCBI Description (AC005508) 9058 [Arabidopsis thaliana] 7783 Seq. No. 10596 5.R1040 Contig ID kl1701203144.h1 5'-most EST BLASTX Method q2894612 NCBI GI BLAST score 170 2.0e-12 E value 57 Match length % identity NCBI Description (AL021889) putative protein [Arabidopsis thaliana] 7784 Seq. No. 10611 1.R1040 Contig ID LIB3039-001-Q1-E1-H1 5'-most EST 7785 Seq. No. 10621 1.R1040 Contig ID LIB3039-001-Q1-E1-F10 5'-most EST Method BLASTX NCBI GI q2769566 204 BLAST score E value 6.0e-16 90 Match length 51 % identity (Y10477) chloroplast thylakoidal processing peptidase NCBI Description [Arabidopsis thaliana] 7786 Seq. No. 10623 1.R1040 Contig ID bth700843521.h1 5'-most EST BLASTX Method q4090257 NCBI GI 291 BLAST score 5.0e-26 E value 91 Match length 63 % identity (AJ131732) ribosomal protein L37A [Pseudotsuga menziesii] NCBI Description 7787 Seq. No. 10623 2.R1040 Contig ID LIB3040-043-Q1-E1-A8

5'-most EST

BLASTX Method q4090257 NCBI GI 473 BLAST score 2.0e-47 E value 92 Match length

1428



% identity NCBI Description (AJ131732) ribosomal protein L37A [Pseudotsuga menziesii] Seq. No. 7788 Contig ID 10623 3.R1040 5'-most EST LIB3087-012-Q1-K1-H12 Method BLASTX NCBI GI g4090257 BLAST score 214 E value 3.0e-17 Match length 42 98 % identity NCBI Description (AJ131732) ribosomal protein L37A [Pseudotsuga menziesii] Seq. No. 7789 Contig ID 10626 1.R1040 5'-most EST g4283326 Method BLASTX NCBI GI g3024516 BLAST score 619 E value 1.0e-97 Match length 217 % identity

RAS-RELATED PROTEIN RAB11C >gi 2160157 (AC000132) Strong NCBI Description

similarity to A. thaliana ara $-\overline{2}$ (gb ATHARA2). ESTs

gb ATTS2483, gb ATTS2484, gb AA042159 come from this gene. [Arabidopsis thaliana] >gi 2231303 (U74669) ras-related

small GTPase [Arabidopsis thaliana]

7790 Seq. No.

Contig ID 10626 2.R1040 5'-most EST pcp700991695.hl

Method BLASTX NCBI GI g1405561 BLAST score 506 E value 2.0e-51 Match length 101 % identity 98

NCBI Description (X98540) FSGTP1 [Fagus sylvatica]

7791 Seq. No.

10638 1.R1040 Contig ID

5'-most EST uC-gmflminsoy072h01b1

Method BLASTX NCBI GI g2459445 BLAST score 815 E value 8.0e-87 323 Match length % identity

NCBI Description (AC002332) putative ribonucleoprotein [Arabidopsis

thaliana]

7792 Seq. No.

Contig ID 10638 2.R1040 5'-most EST ncj700977531.h1

Method BLASTX NCBI GI g2244786



```
196
BLAST score
                   3.0e-15
E value
                  51
Match length
                   37
% identity
NCBI Description (Z97335) ribonucleoprotein homolog [Arabidopsis thaliana]
                   7793
Seq. No.
                  10644 1.R1040
Contig ID
5'-most EST
                  q5607171
                  BLASTX
Method
                   q1707008
NCBI GI
                   547
BLAST score
                   1.0e-77
E value
                   193
Match length
% identity
                   (U78721) 30S ribosomal protein S5 isolog [Arabidopsis
NCBI Description
                   thaliana]
                   7794
Seq. No.
                   10644 2.R1040
Contig ID
                   fjg70\overline{0}968348.h1
5'-most EST
                   BLASTX
Method
                   q1707008
NCBI GI
                   296
BLAST score
                   1.0e-26
E value
                   109
Match length
% identity
                   (U78721) 30S ribosomal protein S5 isolog [Arabidopsis
NCBI Description
                   thaliana]
                   7795
Seq. No.
                   10644 4.R1040
Contig ID
                   6HC-02-Q1-B1-A7
5'-most EST
                   BLASTX
Method
                   g1707008
NCBI GI
BLAST score
                   209
                   1.0e-16
E value
Match length
                   83
% identity
                   (U78721) 30S ribosomal protein S5 isolog [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7796
                   10646 1.R1040
Contig ID
                   jC-qmst02400062c06a1
5'-most EST
                   BLASTN
Method
                   q169980
NCBI GI
                   407
BLAST score
                   0.0e+00
E value
                   920
Match length
 % identity
                  Soybean heat-shock protein (Gmhsp26-A) gene, complete cds
NCBI Description
                   7797
 Seq. No.
                   10646 2.R1040
 Contig ID
```

5'-most EST LIB3094-048-Q1-K1-A5

Method BLASTN

NCBI GI g1184122 BLAST score 346 E value 0.0e+00Match length 650 % identity 88 NCBI Description partial cds

Vigna radiata clone MII-4 auxin-induced protein mRNA,

Seq. No. 7798

10646 3.R1040 Contig ID

5'-most EST LIB3039-001-Q1-E1-E9

Method BLASTN NCBI GI g169980 BLAST score 431 E value 0.0e + 00Match length 1070 % identity 93

NCBI Description Soybean heat-shock protein (Gmhsp26-A) gene, complete cds

7799 Seq. No.

Contig ID 10646 4.R1040 5'-most EST zsq701126407.h1

Method BLASTN NCBI GI q169980 BLAST score 225 E value 1.0e-123 Match length 369 % identity 92

NCBI Description Soybean heat-shock protein (Gmhsp26-A) gene, complete cds

Seq. No. 7800

10646 5.R1040 Contig ID 5'-most EST rlr700900057.h1

Method BLASTN NCBI GI q169980 BLAST score 188 E value 1.0e-101 Match length 509 % identity 88

Soybean heat-shock protein (Gmhsp26-A) gene, complete cds NCBI Description

7801 Seq. No.

10646 6.R1040 Contig ID

5'-most EST LIB3094-051-Q1-K1-A4

BLASTN Method NCBI GI q169980 BLAST score 217 1.0e-118 E value 347 Match length 94 % identity

NCBI Description Soybean heat-shock protein (Gmhsp26-A) gene, complete cds

Seg. No. 7802

Contig ID 10648 1.R1040

5'-most EST LIB3039-001-Q1-E1-C4

Method BLASTN NCBI GI q1518539



```
BLAST score
                  365
                  0.0e + 00
E value
                  765
Match length
                  87
% identity
NCBI Description Glycine max UDP-glucose dehydrogenase mRNA, complete cds
                  7803
Seq. No.
                  10665 1.R1040
Contig ID
                  qsv701051012.h1
5'-most EST
                  BLASTX
Method
                  q2995953
NCBI GI
                   379
BLAST score
                   2.0e-36
E value
                   103
Match length
% identity
                  (AF053565) glutaredoxin I [Mesembryanthemum crystallinum]
NCBI Description
Seq. No.
                   10665 2.R1040
Contig ID
                   bnu700967683.hl
5'-most EST
                   BLASTX
Method
                   q2995953
NCBI GI
                   406
BLAST score
                   2.0e-39
E value
                   105
Match length
% identity
                   (AF053565) glutaredoxin I [Mesembryanthemum crystallinum]
NCBI Description
                   7805
Seq. No.
                   10670 1.R1040
Contig ID
                   leu701156031.h1
5'-most EST
                   BLASTX
Method
                   g481236
NCBI GI
BLAST score
                   279
                   2.0e-24
E value
                   132
Match length
% identity
                   43
                   hypothetical protein - Madagascar periwinkle
NCBI Description
                   >gi_407410_emb_CAA81526_ (Z26880) 14 kDa polypeptide
                   [Catharanthus roseus]
                   7806
Seq. No.
                   10673 1.R1040
 Contig ID
                   vzy700752406.hl
 5'-most EST
                   BLASTX
Method
                   g2136140
NCBI GI
                   793
 BLAST score
                   1.0e-84
 E value
                   272
Match length
 % identity
                   serine C-palmitoyltransferase (EC 2.3.1.50) - human
 NCBI Description
```

(fragment) >gi_1001945 (U15555) serine palmitoyltransferase

[Homo sapiens]

Seq. No. 7807

Contig ID 10677 2.R1040 5'-most EST leu701144384.h1

E value

Match length

6.0e-47

100



```
7808
Seq. No.
                  10682 1.R1040
Contig ID
                  LIB3107-052-Q1-K1-A12
5'-most EST
Method
                  BLASTX
NCBI GI
                  q4490292
                  889
BLAST score
E value
                  8.0e-96
                  218
Match length
% identity
                  (AL035678) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10682 3.R1040
Contig ID
                   jsh70\overline{1}066564.h1
5'-most EST
Method
                   BLASTX
                   q4490292
NCBI GI
BLAST score
                   214
                   2.0e-17
E value
Match length
                   57
% identity
                   (AL035678) putative protein [Arabidopsis thaliana]
NCBI Description
                   7810
Seq. No.
                   10686 1.R1040
Contig ID
5'-most EST
                   jC-gmro02910071f10a1
Method
                   BLASTX
                   g3236238
NCBI GI
BLAST score
                   610
                   2.0e-63
E value
                   129
Match length
% identity
                   86
                   (AC004684) putative ARF1 GTPase activating protein
NCBI Description
                   [Arabidopsis thaliana] >gi_4519792_dbj_BAA75744.1_
                   (AB017876) Asp1 [Arabidopsis thaliana]
                   7811
Seq. No.
                   10686 2.R1040
Contig ID
5'-most EST
                   txt700735901.h1
                   BLASTX
Method
                   g3236238
NCBI GI
BLAST score
                   597
E value
                   4.0e-85
Match length
                   203
                   77
% identity
                   (AC004684) putative ARF1 GTPase activating protein
NCBI Description
                   [Arabidopsis thaliana] >gi 4519792 dbj BAA75744.1
                   (AB017876) Aspl [Arabidopsis thaliana]
Seq. No.
                   7812
Contig ID
                   10686 3.R1040
                   zvj700605112.h2
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3236238
BLAST score
                   468
```



```
% identity
                  (AC004684) putative ARF1 GTPase activating protein
NCBI Description
                  [Arabidopsis thaliana] >gi 4519792 dbj BAA75744.1
                  (AB017876) Asp1 [Arabidopsis thaliana]
Seq. No.
                  7813
                  10687 1.R1040
Contig ID
5'-most EST
                  pmv700892916.h1
                  BLASTX
Method
                  q141435
NCBI GI
BLAST score
                  260
                  2.0e-22
E value
                  119
Match length
                  41
% identity
                  DIHYDRONEOPTERIN ALDOLASE (DHNA) >gi_98360_pir__E37854
NCBI Description
                  folate biosynthesis protein 1 (sul 3 region) - Bacillus
                  subtilis >gi_143411 (M34053) ORF1 [Bacillus subtilis]
                  >gi 467467 dbj BAA05313 (D26185) unknown [Bacillus
                  subtilis] >gi 2632345 emb_CAB11854_ (Z99104)
                  dihydroneopterin aldolase [Bacillus subtilis]
                  7814
Seq. No.
                  10688 1.R1040
Contig ID
                  LIB3039-006-Q1-E1-A1
5'-most EST
                  BLASTX
Method
                  q837
NCBI GI
BLAST score
                  359
                   5.0e-34
E value
                  99
Match length
% identity
                  70
                  (X52113) tryptophan tRNA ligase (AA 1-459) [Bos taurus]
NCBI Description
Seq. No.
                   7815
                   10688 2.R1040
Contia ID
                   uC-gmropic026h04b1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g837
BLAST score
                   349
E value
                   5.0e-33
Match length
                   100
% identity
                   (X52113) tryptophan tRNA ligase (AA 1-459) [Bos taurus]
NCBI Description
                   7816
Seq. No.
                   10697 1.R1040
Contig ID
                   kl1701206518.hl
5'-most EST
                   BLASTX
Method
                   g4006893
NCBI GI
                   1894
BLAST score
                   0.0e + 00
E value
                   517
Match length
% identity
                   68
                   (Z99708) aminopeptidase-like protein [Arabidopsis thaliana]
```

NCBI Description

7817 Seq. No.

Contig ID 10699 1.R1040

5'-most EST LIB3051-097-Q1-K1-B12



```
BLASTN
Method
NCBI GI
                   g456713
BLAST score
                   608
                   0.0e + 00
E value
Match length
                   739
% identity
                   69
NCBI Description Glycine max gene for ubiquitin, complete cds
Seq. No.
                   10699 2.R1040
Contig ID
5'-most EST
                   LIB3040-024-Q1-E1-H12
                   BLASTN
Method
                   g303900
NCBI GI
                   369
BLAST score
                   0.0e+00
E value
Match length
                   586
% identity
                   Soybean gene for ubiquitin, complete cds
NCBI Description
Seq. No.
                   10701 1.R1040
Contig ID
5'-most EST
                   LIB3039-009-Q1-E1-G5
                   BLASTX
Method
                   q2129630
NCBI GI
                   246
BLAST score
                   6.0e-21
E value
Match length
                   91
% identity
                   lamin - Arabidopsis thaliana >gi_1262754 emb_CAA65750
NCBI Description
                   (X97023) lamin [Arabidopsis thaliana] >gi_3395760 (U77721)
                   unknown [Arabidopsis thaliana]
Seq. No.
                   7820
                   10701 2.R1040
Contig ID
                   LIB3109-001-Q1-K1-H8
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2129630
BLAST score
                   150
```

1.0e-09 E value Match length 51 % identity 49

lamin - Arabidopsis thaliana >gi 1262754 emb CAA65750 NCBI Description (X97023) lamin [Arabidopsis thaliana] $>gi_3395760$ (U77721)

unknown [Arabidopsis thaliana]

7821 Seq. No.

10703 1.R1040 Contig ID $leu70\overline{1}146671.h1$ 5'-most EST

BLASTX Method g2760334 NCBI GI BLAST score 392 E value 1.0e-37 90 Match length 76 % identity

(AC002130) F1N21.5 [Arabidopsis thaliana] NCBI Description

Seq. No. 7822

1435



```
10710 1.R1040
Contig ID
                  LIB3039-014-Q1-E1-C6
5'-most EST
                   7823
Seq. No.
                   10718 1.R1040
Contig ID
                   LIB3065-024-Q1-N1-B1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3372233
                   827
BLAST score
                   1.0e-88
E value
                   205
Match length
                   77
% identity
                   (AF019248) RNA polymerase I, II and III 24.3 kDa subunit
NCBI Description
                   [Arabidopsis thaliana]
                   7824
Seq. No.
                   10718 3.R1040
Contig ID
                   jsh70\overline{1}068423.h1
5'-most EST
                   BLASTX
Method
                   q3372233
NCBI GI
                   147
BLAST score
                   2.0e-09
E value
                   30
Match length
                   93
% identity
                   (AF019248) RNA polymerase I, II and III 24.3 kDa subunit
NCBI Description
                   [Arabidopsis thaliana]
                   7825
Seq. No.
                   10733 2.R1040
Contig ID
                   LIB3139-057-P1-N1-B12
5'-most EST
                   7826
Seq. No.
                   10743 1.R1040
Contig ID
                   LIB3087-002-Q1-K1-G10
5'-most EST
                   BLASTX
Method
                   g4406807
NCBI GI
                    691
BLAST score
                    1.0e-72
E value
                    212
Match length
                    66
 % identity
                    (AC006201) putative elongation factor beta-1 [Arabidopsis
NCBI Description
                    thaliana]
                    7827
 Seq. No.
                    10746 1.R1040
 Contig ID
                    LIB3039-030-Q1-E1-H10
 5'-most EST
                    BLASTX
Method
                    g4314355
 NCBI GI
 BLAST score
                    685
                    5.0e-72
 E value
                    207
 Match length
                    58
 % identity
                    (AC006340) unknown protein [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
                    7828
```

10754 1.R1040

LIB3039-036-Q1-E1-A1

Contig ID

5'-most EST

Contig ID

5'-most EST



```
BLASTX
Method
                  g2586123
NCBI GI
                  787
BLAST score
                  6.0e-84
E value
                  192
Match length
                  75
% identity
NCBI Description (U89511) b-keto acyl reductase [Allium porrum]
                  7829
Seq. No.
                  10762 1.R1040
Contig ID
                  LIB3039-038-Q1-E1-E5
5'-most EST
                  BLASTX
Method
                  g1408294
NCBI GI
                  306
BLAST score
                  2.0e-27
E value
                  218
Match length
                   36
% identity
NCBI Description (U61983) benzyl alcohol dehydrogenase [Acinetobacter
                   calcoaceticus]
                   7830
Seq. No.
                   10767 1.R1040
Contig ID
                  LIB3040-052-Q1-E1-D3
5'-most EST
                  BLASTX
Method
                   q4335763
NCBI GI
                   424
BLAST score
                   3.0e-41
E value
                   161
Match length
% identity
NCBI Description (AC006284) unknown protein [Arabidopsis thaliana]
Seq. No.
                   10767 2.R1040
Contig ID
                   fua701040645.hl
5'-most EST
                   7832
Seq. No.
                   10768 1.R1040
Contig ID
                   epx701106195.hl
5'-most EST
                   BLASTX
Method
                   g4098517
NCBI GI
BLAST score
                   684
                   6.0e-72
E value
Match length
                   180
 % identity
                   72
NCBI Description (U79114) auxin-binding protein ABP19 [Prunus persica]
                   7833
 Seq. No.
                   10771 1.R1040
 Contig ID
                   k11701211959.h1
 5'-most EST
                   7834
 Seq. No.
                   10771 2.R1040
 Contig ID
 5'-most EST
                   vzy700754421.h1
                   7835
 Seq. No.
                   10779 1.R1040
```

1437

epx701105157.h1

E value Match length

% identity

NCBI Description

```
Method
                  BLASTX
NCBI GI
                  q119931
BLAST score
                  443
E value
                  1.0e-43
                  146
Match length
                  64
% identity
                  FERREDOXIN I PRECURSOR >gi 65740 pir FEPM1 ferredoxin
NCBI Description
                  [2Fe-2S] I precursor - garden pea >gi 169087 (M31713)
                  ferredoxin I precursor [Pisum sativum]
                  7836
Seq. No.
Contig ID
                  10783 1.R1040
5'-most EST
                  LIB3049-041-Q1-E1-A4
Method
                  BLASTX
NCBI GI
                  g3861188
BLAST score
                  267
                  6.0e-23
E value
Match length
                  103
% identity
                  54
NCBI Description
                  (AJ235272) 50S RIBOSOMAL PROTEIN L24 (rplX) [Rickettsia
                  prowazekii]
Seq. No.
                  7837
                  10783 2.R1040
Contig ID
5'-most EST
                  LIB3049-041-Q1-E1-G6
Method
                  BLASTX
NCBI GI
                  q3861188
BLAST score
                  245
                  1.0e-20
E value
Match length
                  89
% identity
                  53
                  (AJ235272) 50S RIBOSOMAL PROTEIN L24 (rplX) [Rickettsia
NCBI Description
                  prowazekii]
Seq. No.
                  7838
                  10783 3.R1040
Contig ID
5'-most EST
                  LIB3049-039-Q1-E1-A9
Method
                  BLASTX
NCBI GI
                  g143448
BLAST score
                  181
E value
                  2.0e-13
Match length
                  75
                  51
% identity
                  (M81749) ribosomal protein L24 [Bacillus subtilis]
NCBI Description
Seq. No.
                  7839
Contig ID
                  10792 1.R1040
5'-most EST
                  jex70\overline{0}909892.h1
Method
                  BLASTX
NCBI GI
                  g2623298
BLAST score
                  565
```

1438

(AC002409) putative 4-alpha-glucanotransferase [Arabidopsis

3.0e-58

thaliana]

158

```
Seq. No.
                   7840
Contig ID
                   10799 1.R1040
5'-most EST
                   LIB3039-052-Q1-E1-B10
                   7841
Seq. No.
Contig ID
                   10808 1.R1040
5'-most EST
                  LIB3170-018-Q1-K1-D11
                   7842
Seq. No.
                   10809 1.R1040
Contig ID
5'-most EST
                  LIB3040-043-Q1-E1-G12
Method
                  BLASTX
NCBI GI
                   q2462834
BLAST score
                   327
E value
                   6.0e-30
Match length
                  179
% identity
                   41
NCBI Description (AF000657) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  7843
                  10814 1.R1040
Contig ID
5'-most EST
                  LIB3040-061-Q1-E11-H5
Method
                  BLASTX
NCBI GI
                  g2252847
BLAST score
                  162
                   6.0e-11
E value "
Match length
                  62
% identity
                  73
NCBI Description (AF013293) No definition line found [Arabidopsis thaliana]
Seq. No.
                   7844
Contig ID
                   10816 1.R1040
5'-most EST
                   LIB3138-129-Q1-N1-E8
Method
                  BLASTX
NCBI GI
                  g2244847
BLAST score
                  376
E value
                   2.0e-35
Match length
                   531
% identity
                  28
NCBI Description (Z97337) hydroxyproline-rich glycoprotein homolog
                   [Arabidopsis thaliana]
Seq. No.
                  7845
Contig ID
                  10816 2.R1040
5'-most EST
                  uC-gmropic109c08b1
Method
                  BLASTN
NCBI GI
                  q3985934
BLAST score
                  51
E value
                   2.0e-19
Match length
                  211
```

% identity 81

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MJE7, complete sequence [Arabidopsis thaliana]

Seq. No. 7846

Contig ID 10816 4.R1040 5'-most EST kl1701208166.h1

BLAST score

```
7847
Seq. No.
Contig ID
                  10816 5.R1040
5'-most EST
                  LIB3052-007-Q1-B1-G12
Method
                  BLASTN
NCBI GI
                  g3985934
BLAST score
                   44
E value
                   2.0e-15
Match length
                  168
% identity
                   82
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MJE7, complete sequence [Arabidopsis thaliana]
Seq. No.
                  7848
                  10817 1.R1040
Contig ID
                  LIB3040-061-Q1-E11-F8
5'-most EST
Method
                  BLASTN
NCBI GI
                  g3421345
BLAST score
                  194
E value
                  1.0e-105
Match length
                  391
% identity
                   87
                  Orobanche minor chloroplast trnS(GGA) gene, rps4 gene and
NCBI Description
                  trnT(UGU) gene
Seq. No.
                   7849
Contig ID
                   10818 1.R1040
5'-most EST
                   ssr700559718.hl
Seq. No.
                   7850
Contig ID
                   10827 1.R1040
5'-most EST
                  LIB3040-061-Q1-E11-E5
                  BLASTX
Method
NCBI GI
                   g2190419
BLAST score
                   327
E value
                   3.0e - 32
Match length
                  103
% identity
                   69
NCBI Description
                  (Y13632) dem [Lycopersicon esculentum]
                  7851
Seq. No.
                   10829 1.R1040
Contig ID
5'-most EST
                  rlr700900449.hl
Method
                  BLASTX
NCBI GI
                  q3980383
BLAST score
                  298
E value
                   4.0e-27
Match length
                  76
% identity
NCBI Description
                   (AC004561) unknown protein [Arabidopsis thaliana]
                   7852
Seq. No.
                  10829 2.R1040
Contig ID
5'-most EST
                  LIB3093-056-Q1-K1-D6
Method
                  BLASTX
                  g3980383
NCBI GI
```



E value 8.0e-27

Match length 75

% identity 75

NCPL Description (ACCOM561) unknown protein [Arabidonsis

NCBI Description (AC004561) unknown protein [Arabidopsis thaliana]

Seq. No. 7853

Contig ID 10830 1.R1040

5'-most EST LIB3094-017-Q1-K1-A5

Method BLASTN
NCBI GI g3694834
BLAST score 436
E value 0.0e+00
Match length 499
% identity 98

NCBI Description Glycine max alcohol dehydrogenase Adh-1 gene, partial cds

Seq. No. 7854

Contig ID 10830 2.R1040 5'-most EST uC-gmropic091g04b1

Method BLASTN
NCBI GI g452768
BLAST score 319
E value 1.0e-179
Match length 451
% identity 93

NCBI Description P.acutifolius alcohol dehydrogenase-1F mRNA, complete CDS

Seq. No. 7855

Contig ID 10830 3.R1040

5'-most EST LIB3094-022-Q1-K1-C9

Method BLASTN
NCBI GI g3694834
BLAST score 209
E value 1.0e-114
Match length 423
% identity 97

NCBI Description Glycine max alcohol dehydrogenase Adh-1 gene, partial cds

Seq. No. 7856

Contig ID 10830_6.R1040

5'-most EST uC-gmrominsoy039g09b1

Method BLASTN
NCBI GI g4039114
BLAST score 435
E value 0.0e+00
Match length 493
% identity 96

NCBI Description Glycine max alcohol-dehydrogenase (Adh-2) gene, partial cds

Seq. No. 7857

Contig ID 10838 1.R1040

5'-most EST LIB3040-061-Q1-E11-A6

Seq. No. 7858

Contig ID 10838 2.R1040

5'-most EST jC-gmst02400027d03d1



Contig ID 10839_1.R1040

5'-most EST LIB3040-061-Q1-E11-A7

Seq. No. 7860

Contig ID 10845_1.R1040

5'-most EST LIB3040-061-Q1-E11-C2

Method BLASTX
NCBI GI g1709129
BLAST score 1988
E value 0.0e+00
Match length 410
% identity 91

NCBI Description GLYCOGEN SYNTHASE KINASE-3 HOMOLOG MSK-3

>gi_481018_pir__S37642 protein kinase MSK-3 (EC 2.7.1.-) alfalfa >gi_313148_emb_CAA48472_ (X68409) protein kinase

[Medicago sativa]

Seq. No. 7861

Contig ID 10845_3.R1040 5'-most EST eep700865347.h1

Seq. No. 7862

Contig ID 10845_4.R1040 5'-most EST vwf700677426.h1

Seq. No. 7863

Contig ID 10849_2.R1040 5'-most EST leu701150332.h1

Seq. No. 7864

Contig ID 10849_3.R1040 5'-most EST leu701149558.h1

Seq. No. 7865

Contig ID 10853 1.R1040

5'-most EST LIB3040-061-Q1-E1-B9

Seq. No. 7866

Contig ID 10855_1.R1040

5'-most EST LIB3170-020-Q1-J1-D10

Method BLASTX
NCBI GI g3885941
BLAST score 301
E value 3.0e-27
Match length 67
% identity 90

NCBI Description (AF095454) PII protein [Ricinus communis]

Seq. No. 7867

Contig ID 10857_1.R1040

5'-most EST jC-gmst02400029c06a1

Method BLASTX
NCBI GI g82200
BLAST score 533
E value 3.0e-54
Match length 216

% identity 50

NCBI Description hypothetical protein 1244 - common tobacco chloroplast

Seq. No. 7868

Contig ID 10864 1.R1040

5'-most EST LIB3167-023-P4-K4-G3

Method BLASTX
NCBI GI g3600061
BLAST score 262
E value 3.0e-22
Match length 172
% identity 40

NCBI Description (AF080120) contains similarity to DNA binding proteins

[Arabidopsis thaliana]

Seq. No. 7869

Contig ID 10867 1.R1040

5'-most EST LIB3107-035-Q1-K1-G5

Method BLASTX
NCBI GI g3915961
BLAST score 1582
E value 0.0e+00
Match length 559
% identity 88

NCBI Description HYPOTHETICAL 267 KD PROTEIN (ORF 2280)

>gi_2924274_emb_CAA77427_ (Z00044) Ycf2 protein [Nicotiana tabacum] >gi 2924285 emb CAA77438 (Z00044) hypothetical

protein [Nicotiana tabacum]

Seq. No. 7870

Contig ID 10867 2.R1040

5'-most EST LIB3170-005-Q1-K1-E11

Method BLASTN
NCBI GI g2924257
BLAST score 126
E value 3.0e-64
Match length 230
% identity 46

NCBI Description Tobacco chloroplast genome DNA

Seq. No. 7871

Contig ID 10868 1.R1040

5'-most EST uC-gmflminsoy080h03b1

Seq. No. 7872

Contig ID 10873_1.R1040

5'-most EST LIB3040-018-Q1-E1-A11

Method BLASTX
NCBI GI g3582335
BLAST score 844
E value 1.0e-120
Match length 297
% identity 74

NCBI Description (AC005496) unknown protein [Arabidopsis thaliana]

Seq. No. 7873

Contig ID 10873 2.R1040



5'-most EST LIB3040-042-Q1-E1-A10

Seq. No. 7874

10873 3.R1040 Contig ID

5'-most EST LIB3040-036-Q1-E1-D10

7875 Seq. No.

10877 1.R1040 Contig ID

5'-most EST LIB3107-072-Q1-K1-E10

Method BLASTN NCBI GI q343041 BLAST score 167 E value 1.0e-88 Match length 584 % identity 86

NCBI Description Pisum sativum chloroplast Val-tRNA gene

Seq. No. 7876

10878_1.R1040 Contig ID

5'-most EST LIB3040-060-Q1-E1-G9

Method BLASTX NCBI GI g2160183 BLAST score 388 E value 3.0e-42 Match length 118 % identity 79

NCBI Description (ACO00132) Identical to A. thaliana U2 SnRNP-specific A'

protein (gb X69137). ESTs gb ATTS0705, gb ATTS0339 come

from this gene. [Arabidopsis thaliana]

7877 Seq. No.

Contig ID 10888 1.R1040

5'-most EST LIB3040-049-Q1-E1-B7

Method BLASTX NCBI GI g82114 BLAST score 269 E value 2.0e-23 Match length 78 % identity 67

NCBI Description

wound-induced protein - tomato (fragment)
>gi_19320_emb_CAA42537_ (X59882) wound induced protein

[Lycopersicon esculentum]

Seq. No. 7878

Contig ID 10893 1.R1040

5'-most EST LIB3107-030-Q1-K1-B11

Method BLASTX NCBI GI q2760844 BLAST score 350 E value 6.0e-33 Match length 87 79 % identity

NCBI Description (AC003105) hypothetical protein [Arabidopsis thaliana]

7879 Seq. No.

10893 3.R1040 Contig ID 5'-most EST sat701004534.h1

```
5'-most EST
                  LIB3040-060-Q1-E1-E9
                  7881
Seq. No.
                  10902 1.R1040
Contig ID
5'-most EST
                  kl1701211184.h1
Method
                  BLASTX
NCBI GI
                  g2832623
BLAST score
                  631
E value
                  3.0e-65
Match length
                  395
                  37
% identity
                  (AL021711) protein kinase - like protein [Arabidopsis
NCBI Description
                  thaliana]
                  7882
Seq. No.
                  10902 2.R1040
Contig ID
5'-most EST
                  fC-gmro700844396g1
Method
                  BLASTX
NCBI GI
                  q2832623
BLAST score
                  290
E value
                  7.0e-26
Match length
                  135
% identity
                   (AL021711) protein kinase - like protein [Arabidopsis
NCBI Description
                  thaliana]
                   7883
Seq. No.
                  10903 1.R1040
Contig ID
5'-most EST
                  LIB3040-043-Q1-E1-F3
                  BLASTX
Method
NCBI GI
                  g3122724
BLAST score
                   338
E value
                  1.0e-31
Match length
                   69
% identity
                   94
                  60S RIBOSOMAL PROTEIN L38 >gi 2289009 (AC002335) ribosomal
NCBI Description
                  protein L38 isolog [Arabidopsis thaliana]
Seq. No.
                  7884
                  10903 2.R1040
Contig ID
5'-most EST
                  LIB3040-048-Q1-E1-F12
Method
                  BLASTX
NCBI GI
                  q3122724
BLAST score
                   338
E value
                  1.0e-31
Match length
                   69
                  94
% identity
                  60S RIBOSOMAL PROTEIN L38 >gi 2289009 (AC002335) ribosomal
NCBI Description
                  protein L38 isolog [Arabidopsis thaliana]
Seq. No.
                  7885
                  10904 1.R1040
Contig ID
5'-most EST
                  LIB3170-083-Q1-J1-C4
```

7880

10898 1.R1040

Seq. No.

Contig ID



```
7886
Seq. No.
                  10907 1.R1040
Contig ID
5'-most EST
                  LIB3170-019-Q1-J1-D3
                  7887
Seq. No.
                  10912 1.R1040
Contig ID
5'-most EST
                  LIB3170-011-Q1-K1-E9
Seq. No.
                  7888
Contig ID
                  10915 1.R1040
                  zhf700964946.h1
5'-most EST
Method
                  BLASTX
                  g2961378
NCBI GI
BLAST score
                  1482
                  1.0e-165
E value
                  473
Match length
                  37
% identity
NCBI Description
                  (AL022141) putative protein [Arabidopsis thaliana]
Seq. No.
                  10915 2.R1040
Contig ID
5'-most EST
                  uaw700664257.h1
Method
                  BLASTX
NCBI GI
                  g2961378
BLAST score
                  271
E value
                   9.0e-24
Match length
                   69
% identity
                   35
                  (AL022141) putative protein [Arabidopsis thaliana]
NCBI Description
                   7890
Seq. No.
                  10917 1.R1040
Contig ID
5'-most EST
                  uC-gmropic040d12b1
Method
                  BLASTX
NCBI GI
                   q1397319
BLAST score
                   209
                   3.0e-16
E value
Match length
                  118
% identity
                   38
                  (U61953) No definition line found [Caenorhabditis elegans]
NCBI Description
                  7891
Seq. No.
Contig ID
                   10917 2.R1040
5'-most EST
                   LIB3040-060-Q1-E1-C8
Seq. No.
                  7892
Contig ID
                   10920 1.R1040
                  LIB3040-060-Q1-E1-D2
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3193303
```

Method BLASTX
NCBI GI g3193303
BLAST score 234
E value 2.0e-19
Match length 90
% identity 47

NCBI Description (AF069298) similar to several proteins containing a tandem repeat region such as Plasmodium falciparum GGM tandem

repeat protein (GB:U27807); partial CDS [Arabidopsis

thaliana]

Seq. No. 7893

Contig ID 10923_1.R1040

5'-most EST LIB3051-085-Q1-K1-G4

Method BLASTX
NCBI GI g2347194
BLAST score 210
E value 4.0e-18
Match length 182
% identity 36

NCBI Description (AC002338) hypothetical protein [Arabidopsis thaliana]

Seq. No. 7894

Contig ID 10929 1.R1040

5'-most EST LIB3040-052-Q1-E1-H11

Method BLASTX
NCBI GI g3810855
BLAST score 216
E value 5.0e-17
Match length 142
% identity 39

NCBI Description (AL032684) hypothetical protein [Schizosaccharomyces pombe]

Seq. No. 7895

Contig ID 10931 1.R1040

5'-most EST LIB3170-019-Q1-J1-C10

Seq. No. 7896

Contig ID 10935_1.R1040 5'-most EST zpv700762884.h1

Method BLASTX
NCBI GI g1076316
BLAST score 296
E value 1.0e-26
Match length 121
% identity 45

NCBI Description drought-induced protein Dil9 - Arabidopsis thaliana

>gi 469110 emb CAA55321 (X78584) Di19 [Arabidopsis

thaliana]

Seq. No. 7897

Contig ID 10936 1.R1040

5'-most EST LIB3092-056-Q1-K1-G4

Method BLASTX
NCBI GI g3582335
BLAST score 1852
E value 0.0e+00
Match length 375
% identity 91

NCBI Description (AC005496) unknown protein [Arabidopsis thaliana]

Seq. No. 7898

Contig ID 10936 3.R1040

5'-most EST LIB3170-002-Q1-J1-H4

Method BLASTX NCBI GI q3582335

```
BLAST score
                  141
                   1.0e-08
E value
Match length
                   42
                   64
% identity
NCBI Description
                  (AC005496) unknown protein [Arabidopsis thaliana]
                   7899
Seq. No.
Contig ID
                  10936 5.R1040
5'-most EST
                  LIB3072-037-Q1-E1-G5
```

Contig ID 10938 1.R1040

5'-most EST LIB3040-060-Q1-E1-A4

Seq. No. 7901

Contig ID 10941 1.R1040

5'-most EST LIB3040-060-Q1-E1-A9

Seq. No. 7902

Contig ID 10947 1.R1040

5'-most EST LIB3040-060-Q1-E1-B2

Seq. No. 7903

Contig ID 10948_1.R1040 5'-most EST fC-gmse700654525r1

Method BLASTX
NCBI GI g4539423
BLAST score 1549
E value 1.0e-173
Match length 383
% identity 81

NCBI Description (AL049171) pyrophosphate-dependent phosphofructo-1-kinase

[Arabidopsis thaliana]

Seq. No. 7904

Contig ID 10948 2.R1040

5'-most EST LIB3139-008-P1-N1-F6

Seq. No. 7905

Contig ID 10948_3.R1040 5'-most EST fC-gmle700559917a1

Method BLASTX
NCBI GI g4539423
BLAST score 399
E value 2.0e-38
Match length 100
% identity 77

NCBI Description (AL049171) pyrophosphate-dependent phosphofructo-1-kinase

[Arabidopsis thaliana]

Seq. No. 7906

Contig ID 10948_4.R1040

5'-most EST LIB3092-032-Q1-K1-F6

Seq. No. 7907

Contig ID 10948_5.R1040 5'-most EST eep700870309.h1

Contig ID 10954 1.R1040

5'-most EST LIB3051-044-Q1-K1-B3

Method BLASTX
NCBI GI g2583133
BLAST score 247
E value 1.0e-20
Match length 179
% identity 36

NCBI Description (AC002387) unknown protein [Arabidopsis thaliana]

Seq. No. 7909

Contig ID 10954 2.R1040

5'-most EST LIB3109-028-Q1-K1-D4

Seq. No. 7910

Contig ID 10954_3.R1040 5'-most EST leu701148605.h1

Seq. No. 7911

Contig ID 10957 1.R1040 5'-most EST leu701146515.h1

Method BLASTX
NCBI GI g2191136
BLAST score 173
E value 2.0e-16
Match length 176
% identity 34

NCBI Description (AF007269) Similar to UTP-Glucose Glucosyltransferase;

coded for by A. thaliana cDNA T46230; coded for by A. thaliana cDNA H76538; coded for by A. thaliana cDNA H76290

[Arabidopsis thaliana]

Seq. No. 7912

Contig ID 10957 2.R1040

5'-most EST LIB3040-045-Q1-E1-F2

Method BLASTX
NCBI GI g2191136
BLAST score 177
E value 7.0e-13
Match length 100
% identity 37

NCBI Description (AF007269) Similar to UTP-Glucose Glucosyltransferase;

coded for by A. thaliana cDNA T46230; coded for by A. thaliana cDNA H76538; coded for by A. thaliana cDNA H76290

[Arabidopsis thaliana]

Seq. No. 7913

Contig ID 10962 1.R1040

5'-most EST LIB3040-053-Q1-E1-A11

Method BLASTN
NCBI GI g313026
BLAST score 67
E value 4.0e-29
Match length 195
% identity 84

NCBI Description L.esculentum rpl38 mRNA for ribosomal protein L38 7914 Seq. No. 10965 1.R1040 Contig ID LIB3040-059-Q1-E1-D9 5'-most EST 7915 Seq. No. 10972 1.R1040 Contig ID 5'-most EST LIB3106-048-Q1-K1-C3 7916 Seq. No. 10977 1.R1040 Contig ID 5'-most EST LIB3094-038-Q1-K1-H7 7917 Seq. No. 10981 1.R1040 Contig ID 5'-most EST pmv700889551.hl BLASTX Method NCBI GI q2435511 BLAST score 351 5.0e-33 E value Match length 89 % identity 73 (AF024504) contains similarity to prolyl 4-hydroxylase NCBI Description alpha subunit [Arabidopsis thaliana] 7918 Seq. No. 10982 1.R1040 Contig ID 5'-most EST uC-gmropic099a04b1 Method BLASTN g3695058 NCBI GI BLAST score 132 E value 9.0e-68 Match length 344 % identity 90 Lotus japonicus rac GTPase activating protein 1 mRNA, NCBI Description complete cds 7919 Seq. No. 10989_1.R1040 Contig ID 5'-most EST LIB3040-017-Q1-E1-A3 Method BLASTX NCBI GI q2462834

BLAST score 356 2.0e-33 E value Match length 173 % identity 47

(AF000657) hypothetical protein [Arabidopsis thaliana] NCBI Description

7920 Seq. No.

10993 1.R1040 Contig ID leu701154503.h1 5'-most EST

7921 Seq. No.

10994 1.R1040 Contig ID

5'-most EST LIB3040-059-Q1-E1-A8

Method BLASTX

BLAST score

E value

372 2.0e-35

```
NCBI GI
                   g4263713
                   593
BLAST score
E value
                   5.0e-61
Match length
                   200
                   58
% identity
                   (AC006223) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   7922
Seq. No.
Contig ID
                   10994 3.R1040
                   jex70\overline{0}903123.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4263713
BLAST score
                   289
                   4.0e-26
E value
Match length
                   70
% identity
                   71
                   (AC006223) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7923
                   10996 1.R1040
Contig ID
5'-most EST
                  LIB3040-046-Q1-E1-D5
Method
                  BLASTX
NCBI GI
                   q2583135
BLAST score
                   578
                   2.0e-59
E value
Match length
                   159
% identity
                   71
                   (AC002387) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   >gi 3822216 (AF074948) FIL [Arabidopsis thaliana]
                   >gi 4322477 gb AAD16053 (AF087015) abnormal floral organs
                   protein [Arabidopsis thaliana]
                   7924
Seq. No.
Contig ID
                   10996 2.R1040
5'-most EST
                   LIB3170-001-Q1-K1-B1
Method
                   BLASTX
NCBI GI
                   g2583135
BLAST score
                   348
                   7.0e-33
E value
                   115
Match length
                   65
% identity
NCBI Description
                   (AC002387) hypothetical protein [Arabidopsis thaliana]
                   >gi_3822216 (AF074948) FIL [Arabidopsis thaliana]
                   >gi 4322477 gb AAD16053_ (AF087015) abnormal floral organs
                   protein [Arabidopsis thaliana]
Seq. No.
                   7925
                   11001 1.R1040
Contig ID
5'-most EST
                   uC-gmropic062a09b1
                   7926
Seq. No.
                   11002_1.R1040
Contig ID
                   uC-gmflminsoy046e01b1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2464880
```

```
130
Match length
% identity
                   55
                  (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7927
                  11002 2.R1040
Contig ID
5'-most EST
                   fua701042774.h1
                  BLASTX
Method
NCBI GI
                   q2464880
BLAST score
                   292
E value
                   3.0e-26
Match length
                   104
% identity
                   54
NCBI Description
                  (Z99707) putative protein [Arabidopsis thaliana]
                   7928
Seq. No.
                   11004 1.R1040
Contig ID
5'-most EST
                  LIB3170-015-Q1-K1-F6
                   7929
Seq. No.
Contig ID
                  11004 2.R1040
5'-most EST
                  zlv700807692.h1
                   7930
Seq. No.
Contig ID
                   11009 1.R1040
5'-most EST
                  LIB3170-053-Q1-K1-B9
Method
                  BLASTX
                   q4455365
NCBI GI
BLAST score
                   165
                   3.0e-11
E value
Match length
                   35
% identity
                   83
                  (AL035524) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   7931
                   11014 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220081b10a1
                   7932
Seq. No.
                   11016_1.R1040
Contig ID
5'-most EST
                   jC-gmro02910061a02a1
Method
                   BLASTX
                   g2208908
NCBI GI
BLAST score
                   473
E value
                   3.0e-47
Match length
                   108
% identity
                   85
NCBI Description
                   (AB004809) phosphate transporter [Catharanthus roseus]
                   7933
Seq. No.
Contig ID
                   11017 1.R1040
5'-most EST
                   uC-gmropic063b09b1
Method
                   BLASTX
NCBI GI
                   a3461833
BLAST score
                   1056
```

1.0e-115

248

E value Match length % identity

```
(AC004138) putative expansin [Arabidopsis thaliana]
NCBI Description
                  7934
Seq. No.
                  11017_2.R1040
Contig ID
5'-most EST
                  fC-gmse700658228z1
Method
                  BLASTX
NCBI GI
                   q4027897
BLAST score
                  268
E value
                   4.0e-23
Match length
                  55
% identity
                   84
                  (AF049353) alpha-expansin precursor [Nicotiana tabacum]
NCBI Description
                   7935
Seq. No.
                  11017 3.R1040
Contig ID
                  LIB3040-058-Q1-E1-F4
5'-most EST
                  BLASTX
Method
NCBI GI
                   g4027897
BLAST score
                   706
                   1.0e-74
E value
Match length
                  162
                   79
% identity
                   (AF049353) alpha-expansin precursor [Nicotiana tabacum]
NCBI Description
Seq. No.
                   7936
                   11018 1.R1040
Contig ID
                  LIB3040-060-Q1-E1-A3
5'-most EST
                  BLASTX
Method
NCBI GI
                   q2253442
BLAST score
                   260
                   3.0e-22
E value
Match length
                   62
% identity
                   66
NCBI Description
                   (AF007784) LTCOR11 [Lavatera thuringiaca]
                   7937
Seq. No.
                   11019 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy277b05b1
Method
                   BLASTX
                   g4158232
NCBI GI
BLAST score
                   1716
E value
                   0.0e + 00
Match length
                   332
% identity
NCBI Description
                   (Y18626) reversibly glycosylated polypeptide [Triticum
                   aestivum]
Seq. No.
                   7938
Contig ID
                   11019 2.R1040
5'-most EST
                   kl1701204287.h2
Method
                   BLASTN
                   g2218151
NCBI GI
BLAST score
                   228
E value
                   1.0e-125
Match length
                   404
% identity
                   89
```

NCBI Description Vigna unguiculata type IIIa membrane protein cp-wap13 mRNA, complete cds 7939 Seq. No. 11019 3.R1040 Contig ID 5'-most EST jC-gmst02400003d10d1 BLASTN Method NCBI GI g2218151 BLAST score 137 6.0e-71 E value Match length 259 % identity 92 Vigna unquiculata type IIIa membrane protein cp-wap13 mRNA, NCBI Description complete cds 7940 Seq. No. 11029 1.R1040 Contig ID g5666720 5'-most EST Method BLASTX NCBI GI q2495209 BLAST score 382 E value 2.0e-36 Match length 105 % identity 62 HYPOTHETICAL 12.9 KD PROTEIN SLR1417 NCBI Description >gi 1651819 dbj BAA16746 (D90900) hypothetical protein [Synechocystis sp.] 7941 Seq. No. Contig ID 11029 2.R1040 leu701148522.h1 5'-most EST 7942 Seq. No. 11030 1.R1040 Contig ID 5'-most EST wrg700788995.h2 BLASTX Method NCBI GI q4263519 BLAST score 469 6.0e-47 E value Match length 96 % identity 94 (AC004044) small nuclear riboprotein Sm-D1 [Arabidopsis NCBI Description thaliana] 7943 Seq. No. Contig ID 11030 2.R1040 5'-most EST LIB3040-058-Q1-E1-G9 Method BLASTX

g4263519

NCBI GI BLAST score 393 5.0e-38 E value Match length 79 % identity 97

(AC004044) small nuclear riboprotein Sm-D1 [Arabidopsis NCBI Description

thaliana]

Seq. No. 7944

```
Contig ID
                  11035 1.R1040
5'-most EST
                  LIB3106-058-Q1-K1-A9
                  BLASTN
Method
NCBI GI
                  g3821780
BLAST score
                  36
                  1.0e-10
E value
Match length
                  36
                  100
% identity
NCBI Description Xenopus laevis cDNA clone 27A6-1
Seq. No.
                  7945
                  11037 1.R1040
Contig ID
5'-most EST
                  hyd700727479.h1
Method
                  BLASTX
NCBI GI
                  g4567203
BLAST score
                  426
E value
                  1.0e-41
                  87
Match length
% identity
                  91
                  (AC007168) putative beta-hydroxyacyl-ACP dehydratase
NCBI Description
                  [Arabidopsis thaliana]
                  7946
Seq. No.
Contig ID
                  11044 1.R1040
5'-most EST
                  sat701015360.h1
                  7947
Seq. No.
                  11044 2.R1040
Contig ID
5'-most EST
                  LIB3055-003-Q1-N1-G2
                  7948
Seq. No.
Contig ID
                  11045 1.R1040
5'-most EST
                  crh700850109.h1
Method
                  BLASTX
NCBI GI
                  g1173624
BLAST score
                  1102
E value
                  1.0e-120
                  342
Match length
% identity
                  63
                  (U34744) cytochrome P-450 [Phalaenopsis sp. 'hybrid
NCBI Description
                  SM9108']
                  7949
Seq. No.
                  11045 2.R1040
Contig ID
5'-most EST
                  LIB3170-020-Q1-K1-B5
Method
                  BLASTX
NCBI GI
                  q1173624
BLAST score
                  259
                  3.0e-22
E value
Match length
                  77
% identity
                  (U34744) cytochrome P-450 [Phalaenopsis sp. 'hybrid
NCBI Description
                  SM9108']
```

Contig ID 11045 4.R1040 5'-most EST vwf700674610.h1



7951 Seq. No. 11053 1.R1040 Contig ID 5'-most EST $g5753\overline{1}61$ 7952 Seq. No. 11060 1.R1040 Contig ID LIB3072-053-Q1-E1-G11 5'-most EST BLASTX Method q126078 NCBI GI BLAST score 277 2.0e-24 E value 117 Match length 30 % identity NCBI Description LATE EMBRYOGENESIS ABUNDANT PROTEIN D-34 (LEA D-34) >gi 81554 pir S04046 embryonic abundant protein gD-34 upland cotton >gi_18501_emb_CAA31594_ (X13206) D-34 Lea protein [Gossypium hirsutum] >gi_167385 (M19389) storage protein [Gossypium hirsutum] >gi_226556_prf__1601521F Lea D-34 gene [Saguinus oedipus] 7953 Seq. No. 11061 1.R1040 Contig ID LIB3040-044-Q1-E1-A3 5'-most EST BLASTX Method q2244822 NCBI GI BLAST score 328 1.0e-30 E value 85 Match length 72 % identity (Z97336) RNA polymerase II fifth largest subunit homolog NCBI Description [Arabidopsis thaliana] 7954 Seq. No. 11063 1.R1040 Contig ID LIB3092-008-Q1-K1-D1 5'-most EST Method BLASTN g4104973 NCBI GI 491 BLAST score 0.0e+00E value Match length 895 89 % identity Pisum sativum spermidine synthase 2 (SPDSYN2) mRNA, NCBI Description complete cds 7955 Seq. No. 11063_2.R1040 Contig ID LIB3052-014-Q1-N1-H2 5'-most EST BLASTN Method g4104971 NCBI GI 205 BLAST score 1.0e-111 E value Match length 347

Pisum sativum spermidine synthase 1 (SPDSYN1) mRNA,

90

complete cds

% identity

NCBI Description

```
Seq. No.
                   7956
Contig ID
                   11063 3.R1040
5'-most EST
                   asn701133255.h1
Method
                   BLASTN
NCBI GI
                   g4104973
BLAST score
                   34
E value
                   1.0e-09
Match length
                   50
% identity
                   92
NCBI Description
                  Pisum sativum spermidine synthase 2 (SPDSYN2) mRNA,
                   complete cds
Seq. No.
                   7957
Contig ID
                   11063 4.R1040
5'-most EST
                   jC-gmle01810063d11d1
Method
                   BLASTN
NCBI GI
                   g4104973
BLAST score
                   66
E value
                   1.0e-28
Match length
                   193
% identity
                   89
NCBI Description
                  Pisum sativum spermidine synthase 2 (SPDSYN2) mRNA,
                  complete cds
Seq. No.
                   7958
Contig ID
                   11063 5.R1040
5'-most EST
                   g4260078
Method
                  BLASTN
NCBI GI
                  g4104971
BLAST score
                  120
E value
                   8.0e-61
Match length
                  245
% identity
                   90
NCBI Description
                  Pisum sativum spermidine synthase 1 (SPDSYN1) mRNA,
                  complete cds
Seq. No.
                  7959
Contig ID
                  11065 1.R1040
5'-most EST
                  jC-qmle01810068a04a1
Method
                  BLASTX
NCBI GI
                  q3805842
BLAST score
                  285
E value
                  3.0e-25
Match length
                  161
% identity
NCBI Description
                  (AL031986) putative protein [Arabidopsis thaliana]
Seq. No.
                  7960
Contig ID
                  11066 1.R1040
5'-most EST
                  LIB3040-057-Q1-E1-H1
```

Method BLASTX NCBI GI g2119045 BLAST score 454 E value 3.0e-45Match length 99 % identity 87

NCBI Description small nuclear ribonucleoprotein U1A - potato



>gi_1050840_emb_CAA90282_ (Z49990) UlsnRNP-specific
protein, UlA [Solanum tuberosum]

Seq. No. 7961

Contig ID 11069_1.R1040

5'-most EST jC-gmro02910052d01a1

Method BLASTX
NCBI GI 9537319
BLAST score 1386
E value 1.0e-154
Match length 309
% identity 83

NCBI Description (L36158) peroxidase [Medicago sativa]

Seq. No. 7962

Contig ID 11072 1.R1040

5'-most EST LIB3049-056-Q1-E1-H3

Method BLASTX
NCBI GI g3881724
BLAST score 291
E value 7.0e-26
Match length 90
% identity 58

NCBI Description (Z69385) Similarity to Yeast JTA107 protein (PIR Acc. No.

S55137); cDNA EST yk290e3.3 comes from this gene; cDNA EST yk290e3.5 comes from this gene [Caenorhabditis elegans]

Seq. No. 7963

Contig ID 11072 2.R1040

5'-most EST LIB3040-057-Q1-E1-H9

Method BLASTX
NCBI GI g3881724
BLAST score 272
E value 5.0e-24
Match length 89
% identity 55

NCBI Description (Z69385) Similarity to Yeast JTA107 protein (PIR Acc. No. S55137); cDNA EST yk290e3.3 comes from this gene; cDNA EST

yk290e3.5 comes from this gene [Caenorhabditis elegans]

Seq. No. 7964

Contig ID 11074 1.R1040 5'-most EST sat701014926.h1

Method BLASTX
NCBI GI g3834307
BLAST score 341
E value 7.0e-32
Match length 94
% identity 72

NCBI Description (AC005679) Strong similarity to gene T10I14.120 gi 2832679

putative protein from Arabidopsis thaliana BAC gb AL021712.

ESTs gb_N65887 and gb_N65627 come from this gene.

[Arabidopsis thaliana]

Seq. No. 7965

Contig ID 11075 1.R1040

5'-most EST LIB3170-017-Q1-J1-A11

```
7966
Seq. No.
Contig ID
                  11077 1.R1040
5'-most EST
                  LIB3170-017-Q1-J1-A12
                  7967
Seq. No.
                  11082 1.R1040
Contig ID
5'-most EST
                  LIB3040-057-Q1-E1-E9
                  7968
Seq. No.
                  11083 1.R1040
Contig ID
5'-most EST
                  LIB3040-031-Q1-E2-C12
Method
                  BLASTX
NCBI GI
                  g3283409
BLAST score
                  165
                  4.0e-11
E value
Match length
                  58
                  59
% identity
                  (AF068754) heat shock factor binding protein 1 HSBP1 [Homo
NCBI Description
                  sapiens] >gi 4557647 ref NP 001528.1 pHSBP1 heat shock
                  factor binding protein
                  7969
Seq. No.
Contig ID
                  11085 1.R1040
5'-most EST
                  LIB3074-036-Q1-K1-E2
Method
                  BLASTX
NCBI GI
                  g2739375
BLAST score
                  268
                  2.0e-23
E value
Match length
                  107
% identity
                  55
NCBI Description
                  (AC002505) unknown protein [Arabidopsis thaliana]
Seq. No.
                  7970
Contig ID
                  11085 3.R1040
5'-most EST
                  LIB3106-094-Q1-K1-F3
Method
                  BLASTX
NCBI GI
                  q2739375
BLAST score
                  152
E value
                  4.0e-10
Match length
                  46
                  67
% identity
NCBI Description
                  (AC002505) unknown protein [Arabidopsis thaliana]
Seq. No.
                  7971
Contig ID
                  11093 1.R1040
```

5'-most EST LIB3106-057-Q1-K1-G9

7972

Method BLASTN NCBI GI g3603400 BLAST score 455 E value 0.0e+00Match length 1073 % identity

NCBI Description Medicago sativa cinnamyl-alcohol dehydrogenase (MsaCadl)

1459

mRNA, complete cds

Seq. No.

Contig ID

7978

11109_1.R1040



```
11093 3.R1040
Contig ID
                  zhf700954121.h1
5'-most EST
Method
                  BLASTN
NCBI GI
                  q556421
BLAST score
                  114
E value
                  3.0e-57
                  274
Match length
                  85
% identity
                  Stylosanthes humilis cinnamyl alcohol dehydrogenase (CAD1)
NCBI Description
                  mRNA, complete cds
                  7973
Seq. No.
                   11095 1.R1040
Contig ID
5'-most EST
                  LIB3040-057-Q1-E1-G2
Seq. No.
                   7974
Contig ID
                   11100 1.R1040
5'-most EST
                  LIB3170-018-Q1-J1-A3
Seq. No.
                   7975
Contig ID
                   11102 1.R1040
5'-most EST
                  LIB3107-042-Q1-K1-C6
Method
                   BLASTX
NCBI GI
                   q3738328
BLAST score
                   315
E value
                   5.0e-29
Match length
                   79
% identity
                   72
                   (AC005170) putative serine carboxypeptidase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   7976
                   11102 2.R1040
Contig ID
                   LIB3040-057-Q1-E1-D3
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4263777
BLAST score
                   234
                   1.0e-19
E value
Match length
                   84
% identity
                   55
                   (AC006068) putative serine carboxypeptidase II [Arabidopsis
NCBI Description
                   thaliana] >gi 4510391 gb AAD21479.1 (AC007017) putative
                   serine carboxypeptidase II [Arabidopsis thaliana]
                   7977
Seq. No.
Contig ID
                   11103 1.R1040
5'-most EST
                   LIB3040-057-Q1-E1-D5
Method
                   BLASTX
                   g3004550
NCBI GI
BLAST score
                   176
E value
                   2.0e-12
Match length
                   64
                   50
% identity
                   (AC003673) hypothetical protein [Arabidopsis thaliana]
NCBI Description
```

```
LIB3040-057-Q1-E1-E11
5'-most EST
                  BLASTX
Method
                  q586079
NCBI GI
                  1015
BLAST score
                  1.0e-111
E value
Match length
                  196
% identity
                  TUBULIN GAMMA-1 CHAIN >gi 460089 (U02069) gl-tubulin
NCBI Description
                  [Arabidopsis thaliana]
                  7979
Seq. No.
                  11114 1.R1040
Contig ID
                  LIB3040-057-Q1-E1-B1
5'-most EST
                  7980
Seq. No.
                  11120 1.R1040
Contig ID
5'-most EST
                  LIB3107-059-Q1-K1-D10
                  BLASTX
Method
                  q266410
NCBI GI
                  378
BLAST score
                   4.0e-36
E value
                   144
Match length
                   62
% identity
                   CDC2+/CDC28-RELATED PROTEIN KINASE R2 >gi_82507_pir__S13934
NCBI Description
                   protein kinase (EC 2.7.1.37) chain cdc2/cdc28 homolog -
                   rice >gi_20194 emb_CAA41172_ (X58194) cdc2+/CDC28-related
                   protein kinase [Oryza sativa]
Seq. No.
                   11122 1.R1040
Contig ID
                   LIB3170-017-Q1-J1-A1
5'-most EST
                   7982
Seq. No.
                   11126 1.R1040
Contig ID
                   LIB3170-020-Q1-K1-A12
5'-most EST
                   7983
Seq. No.
                   11127 1.R1040
Contig ID
                   LIB3040-057-Q1-E1-C4
 5'-most EST
                   BLASTX
Method
                   q4508075
NCBI GI
BLAST score
                   639
                   2.0e-66
E value
                   339
Match length
                   44
 % identity
                   (AC005882) 50259 [Arabidopsis thaliana]
 NCBI Description
                   7984
 Seq. No.
                   11128 1.R1040
 Contig ID
                   LIB3139-062-P1-N1-H1
 5'-most EST
                   BLASTX
 Method
 NCBI GI
                   q2388956
                   629
 BLAST score
                   2.0e-65
 E value
                   223
 Match length
```

1461

58

NCBI Description (Z98979) ribulose-phosphate 3-epimerase

% identity



[Schizosaccharomyces pombe]

Seq. No. 7985

Contig ID 11128_2.R1040

5'-most EST LIB3139-119-P1-N1-D7

Method BLASTX
NCBI GI g861366
BLAST score 316
E value 4.0e-29
Match length 121
% identity 54

NCBI Description (U28991) coded for by C. elegans cDNA cm21c7

[Caenorhabditis elegans]

Seq. No. 7986

Contig ID 11129 1.R1040

5'-most EST LIB3040-057-Q1-E1-C7

Seq. No. 7987

Contig ID 11134 1.R1040

5'-most EST LIB3170-019-Q1-K1-A5

Seq. No. 7988

Contig ID 11143 1.R1040

5'-most EST LIB3049-029-Q1-E1-A1

Seq. No. 7989

Contig ID 11145 1.R1040

5'-most EST LIB3107-002-Q1-K1-E3

Method BLASTX
NCBI GI g2388575
BLAST score 233
E value 2.0e-19
Match length 75

% identity 53

NCBI Description (AC000098) YUP8H12.18 [Arabidopsis thaliana]

Seq. No. 7990

Contig ID 11151_1.R1040

5'-most EST LIB3040-056-Q1-E1-F6

Seq. No. 7991

Contig ID 11156_1.R1040

5'-most EST LIB3170-020-Q1-K1-A5

Seq. No. 7992

Contig ID 11161_1.R1040

5'-most EST LIB3052-007-Q1-B1-E5

Seq. No. 7993

Contig ID 11168_1.R1040

5'-most EST LIB3040-056-Q1-E1-D8

Seq. No. 7994

Contig ID 11172 1.R1040

5'-most EST LIB3049-008-Q1-E1-A6



```
7995
Seq. No.
                  11172 2.R1040
Contig ID
                  LIB3040-056-Q1-E1-C5
5'-most EST
                  BLASTX
Method
                   g4559384
NCBI GI
                   215
BLAST score
E value
                   3.0e-17
                   62
Match length
% identity
                  (AC006526) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7996
Seq. No.
                   11172 3.R1040
Contig ID
                   LIB30\overline{5}1-108-Q1-K1-F6
5'-most EST
                   BLASTX
Method
                   g4559384
NCBI GI
BLAST score
                   164
                   5.0e-11
E value
                   65
Match length
                   55
% identity
                  (AC006526) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7997
Seq. No.
                   11183 1.R1040
Contig ID
                   LIB3050-007-Q1-E1-H10
5'-most EST
Seq. No.
                   7998
                   11184 1.R1040
Contig ID
                   leu701152062.h1
5'-most EST
Method
                   BLASTX
                   q3785978
NCBI GI
                   444
BLAST score
E value
                   5.0e-44
                   132
Match length
% identity
                   (AC005560) unknown protein [Arabidopsis thaliana]
NCBI Description
                   7999
Seq. No.
                   11185 1.R1040
Contig ID
                   LIB3040-056-Q1-E1-B4
5'-most EST
Seq. No.
                   8000
                   11190 1.R1040
Contig ID
                   LIB3040-056-Q1-E1-C10
5'-most EST
                   BLASTX
Method
                   g2980795
NCBI GI
                   825
BLAST score
                   2.0e-88
E value
                   177
Match length
                   89
% identity
                   (AL022197) putative protein [Arabidopsis thaliana]
NCBI Description
                   8001
Seq. No.
```

Contig ID 11192 1.R1040 5'-most EST epx701104273.h1

Method BLASTX NCBI GI g4115916



```
BLAST score
                  145
                  3.0e-09
E value
Match length
                  96
                   46
% identity
                   (AF118222) F3H7.9 gene product [Arabidopsis thaliana]
NCBI Description
                  >gi 4539441_emb_CAB40029.1_ (AL049523) putative protein
                   [Arabidopsis thaliana]
                  8002
Seq. No.
Contig ID
                  11195 1.R1040
                  LIB3040-056-Q1-E1-A5
5'-most EST
                   8003
Seq. No.
                   11195 2.R1040
Contig ID
5'-most EST
                   gsv701051708.hl
Seq. No.
                   8004
Contig ID
                   11198 1.R1040
5'-most EST
                   LIB3040-055-Q1-E1-G6
Method
                   BLASTX
NCBI GI
                   q2208988
BLAST score
                   396
                   9.0e-39
E value
Match length
                   88
% identity
                  (Y10117) signal recognition particle subunit 9 [Zea mays]
NCBI Description
                   8005
Seq. No.
                   11204 1.R1040
Contig ID
                   uC-qmflminsoy058d10b1
5'-most EST
                   BLASTN
Method
                   g294665
NCBI GI
BLAST score
                   216
                   1.0e-117
E value
                   512
Match length
                   86
% identity
                   Castor bean chloroplast beta-ketoacyl-ACP synthase mRNA,
NCBI Description
                   complete cds
                   8006
Seq. No.
                   11211 1.R1040
Contig ID
                   LIB3040-056-Q1-E1-A4
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4432814
BLAST score
                   408
                   8.0e-40
E value
                   92
Match length
```

84 % identity

NCBI Description (AC006593) unknown protein [Arabidopsis thaliana]

Seq. No. 8007

11211_2.R1040 Contig ID uC-qmronoir007c10b1 5'-most EST

BLASTX Method q4432814 NCBI GI BLAST score 219 E value 8.0e-18

```
Match length
                  50
                  80
% identity
                  (AC006593) unknown protein [Arabidopsis thaliana]
NCBI Description
                  8008
Seq. No.
Contig ID
                  11217 1.R1040
                  sat701010626.hl
5'-most EST
                  BLASTX
Method
                  q3786001
NCBI GI
BLAST score
                  478
E value
                  2.0e-85
                  200
Match length
% identity
                  (AC005499) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  8009
                  11218 1.R1040
Contig ID
5'-most EST
                  LIB3040-055-Q1-E1-F10
                  BLASTX
Method
                  q2335098
NCBI GI
BLAST score
                  1103
E value
                  1.0e-121
                   331
Match length
% identity
                  (AC002339) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   8010
Contig ID
                   11223 1.R1040
                   LIB3040-055-Q1-E1-F7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4165861
BLAST score
                   187
                   6.0e-14
E value
Match length
                   83
% identity
                   46
                  (AF006603) histone deacetylase mHDA2 [Mus musculus]
NCBI Description
                   8011
Seq. No.
                   11231 1.R1040
Contig ID
                   LIB3040-055-Q1-E1-C3
5'-most EST
Seq. No.
                   8012
                   11233 1.R1040
Contig ID
                   jC-gmro02910068a07a1
5'-most EST
                   BLASTX
Method
                   g4049341
NCBI GI
                   204
BLAST score
                   9.0e-16
E value
                   71
Match length
                   54
% identity
                  (AL034567) putative protein [Arabidopsis thaliana]
```

NCBI Description

11235 1.R1040 Contig ID

LIB3040-055-Q1-E1-D1 5'-most EST

8014 Seq. No.



Contig ID 11237_1.R1040

5'-most EST LIB3050-008-Q1-E1-F4

Method BLASTX
NCBI GI g461530
BLAST score 454
E value 6.0e-45
Match length 173
% identity 55

NCBI Description N-TERMINAL ACETYLTRANSFERASE COMPLEX ARD1 SUBUNIT HOMOLOG

>qi 159365 (L11348) ARD1 protein homologue [Leishmania

donovani]

Seq. No. 8015

Contig ID 11240 1.R1040

5'-most EST LIB3170-013-Q1-K1-H4

Method BLASTX
NCBI GI g3184082
BLAST score 407
E value 2.0e-39
Match length 156
% identity 47

NCBI Description (AL023781) N-terminal acetyltransferase 1

[Schizosaccharomyces pombe]

Seq. No. 8016

Contig ID 11242_1.R1040

5'-most EST LIB3139-026-P1-N1-F7

Method BLASTX
NCBI GI g3549670
BLAST score 1829
E value 0.0e+00
Match length 424
% identity 80

NCBI Description (AL031394) putative protein [Arabidopsis thaliana]

Seq. No. 8017

Contig ID 11243 1.R1040

5'-most EST LIB31 $\overline{7}0-014-Q1-J1-H2$

Seq. No. 8018

Contig ID 11245 1.R1040

5'-most EST LIB3040-046-Q1-E1-H9

Seq. No. 8019

Contig ID 11258 1.R1040

5'-most EST LIB3040-055-Q1-E1-B12

Method BLASTX
NCBI GI 94539355
BLAST score 152
E value 8.0e-10
Match length 53
% identity 60

NCBI Description (AL049525) putative protein [Arabidopsis thaliana]

Seq. No. 8020

Contig ID 11270_1.R1040 5'-most EST g4291877



Method BLASTX
NCBI GI g1532167
BLAST score 200
E value 3.0e-15
Match length 65
% identity 51

NCBI Description (U63815) localized according to blastn similarity to EST sequences; therefore, the coding span corresponds only to an area of similarity since the initation codon and stop codon could not be precisely determined [Arabidopsis

thaliana]

Seq. No. 8021

Contig ID 11278 1.R1040

5'-most EST LIB3107-060-Q1-K1-D12

Method BLASTX
NCBI GI g3831452
BLAST score 508
E value 2.0e-51
Match length 135
% identity 67

NCBI Description (AC005700) putative cytochrome P450 [Arabidopsis thaliana]

Seq. No. 8022

Contig ID 11278_2.R1040 5'-most EST zsg701127707.h1

Method BLASTX
NCBI GI g3831452
BLAST score 423
E value 1.0e-41
Match length 111
% identity 64

NCBI Description (AC005700) putative cytochrome P450 [Arabidopsis thaliana]

Seq. No. 8023

Contig ID 11283_1.R1040

5'-most EST LIB3073-019-Q1-K1-A7

Method BLASTX
NCBI GI g3184288
BLAST score 223
E value 5.0e-18
Match length 60
% identity 70

NCBI Description (AC004136) unknown protein [Arabidopsis thaliana]

Seq. No. 8024

Contig ID 11283_4.R1040 5'-most EST vwf700675923.h1

Method BLASTX
NCBI GI g3184288
BLAST score 210
E value 5.0e-17
Match length 50
% identity 74

NCBI Description (AC004136) unknown protein [Arabidopsis thaliana]

Seq. No. 8025



```
Contig ID 11287_1.R1040
5'-most EST uC-gmronoir032c05b1
Method BLASTX
NCBI GI g2894598
BLAST score 441
E value 1.0e-43
Match length 93
```

% identity 85 NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8026

Contig ID 11287_2.R1040 5'-most EST sat701010684.h1

Method BLASTX
NCBI GI g2894598
BLAST score 387
E value 2.0e-37
Match length 81
% identity 85

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8027

Contig ID 11303_1.R1040 5'-most EST pcp700989501.h1

Seq. No. 8028

Contig ID 11309_1.R1040

5'-most EST LIB3040-053-Q1-E1-G7

Method BLASTX
NCBI GI g1439609
BLAST score 608
E value 4.0e-63
Match length 164
% identity 46

NCBI Description (U62778) delta-tonoplast intrinsic protein [Gossypium

hirsutum]

Seq. No. 8029

Contig ID 11310_1.R1040 5'-most EST asn701139329.h1

Method BLASTN
NCBI GI g2924257
BLAST score 114
E value 6.0e-57
Match length 394
% identity 90

NCBI Description Tobacco chloroplast genome DNA

Seq. No. 8030

Contig ID 11312 1.R1040

5'-most EST LIB3170-013-Q1-K1-G5

Seq. No. 8031

Contig ID 11316_1.R1040

5'-most EST LIB3040-053-Q1-E1-H9

Seq. No. 8032



Contig ID 11318 1.R1040 5'-most EST LIB3072-050-Q1-K1-E12 Seq. No. 8033 Contig ID 11319 1.R1040 5'-most EST LIB3073-004-Q1-K1-G12 Method BLASTX NCBI GI q2281089 BLAST score 408 E value 1.0e-39 Match length 91 % identity 86 NCBI Description (AC002333) Sm protein F isolog [Arabidopsis thaliana] Seq. No. 8034 11322 1.R1040 Contig ID 5'-most EST LIB3040-053-Q1-E1-D9 8035 Seq. No. Contig ID 11326 1.R1040 5'-most EST LIB3040-053-Q1-E1-E2 Method BLASTX NCBI GI g3650037 BLAST score 989 E value 1.0e-107 Match length 214 % identity 89 (AC005396) auxin-responsive GH3-like protein [Arabidopsis NCBI Description thaliana] 8036 Seq. No. 11329 1.R1040 Contig ID 5'-most EST LIB3049-017-Q1-E1-A7 Method BLASTX NCBI GI g4263771 BLAST score 298 E value 5.0e-27 Match length 86 % identity 57 NCBI Description (AC006218) putative nonspecific lipid-transfer protein precursor [Arabidopsis thaliana] Seq. No. 8037 Contig ID 11333 1.R1040 5'-most EST LIB3094-079-Q1-K1-F12 Method BLASTX NCBI GI g4538911 BLAST score 398 E value 3.0e-38 Match length 175

% identity 50

NCBI Description (AL049482) hypothetical protein [Arabidopsis thaliana]

8038 Seq. No.

11333 2.R1040 Contig ID

5'-most EST LIB3107-076-Q1-K1-C2



Contig ID 11336_1.R1040

5'-most EST LIB3051-047-Q1-K1-D3

Seq. No. 8040

Contig ID 11336_2.R1040

5'-most EST uC-gmflminsoy008d12b1

Seq. No. 8041

Contig ID 11336 3.R1040

5'-most EST LIB3040-045-Q1-E1-C3

Seq. No. 8042

Contig ID 11341 1.R1040

5'-most EST LIB3040-053-Q1-E1-B7

Seq. No. 8043

Contig ID 11343 1.R1040 5'-most EST kl1701206623.h1

Seq. No. 8044

Contig ID 11343_2.R1040 5'-most EST ncj700983657.h1

Seq. No. 8045

Contig ID 11344 1.R1040

5'-most EST jC-gmf102220100e03a1

Seq. No. 8046

Contig ID 11348 1.R1040

5'-most EST LIB3040-053-Q1-E1-D11

Seq. No. 8047

Contig ID 11356_1.R1040

5'-most EST LIB3040-052-Q1-E1-H6

Method BLASTN
NCBI GI g2244991
BLAST score 44

E value 1.0e-15
Match length 112
% identity 85

NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I contig

fragment No

Seq. No. 8048

Contig ID 11363_1.R1040

5'-most EST LIB3139-050-P1-N1-F10

Method BLASTX
NCBI GI g4220477
BLAST score 383
E value 6.0e-37
Match length 154
% identity 58

NCBI Description (AC006069) unknown protein [Arabidopsis thaliana]

Seq. No. 8049

Contig ID 11365_1.R1040



```
rlr700896481.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                   q2244846
BLAST score
                   466
E value
                   2.0e-46
Match length
                   102
% identity
                  (Z97337) ferredoxin [Arabidopsis thaliana]
NCBI Description
Seq. No.
Contig ID
                   11367 1.R1040
                   seb70\overline{0}650601.h1
5'-most EST
                   8051
Seq. No.
Contig ID
                   11370 1.R1040
5'-most EST
                   LIB3138-062-Q1-N1-E3
Seq. No.
                   11372 1.R1040
Contig ID
                   awf700840180.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   a1172995
BLAST score
                   272
                   1.0e-23
E value
                   107
Match length
                   54
% identity
                   60S RIBOSOMAL PROTEIN L22 >gi 1083790 pir S52084 ribosomal
NCBI Description
                   protein L22 - rat >gi 710295 emb_CAA55204_ (X78444)
                   ribosomal protein L22 [Rattus norvegicus]
                   >gi_1093952_prf__2105193A ribosomal protein L22 [Rattus
                   norvegicus]
Seq. No.
                   8053
Contig ID
                   11378 1.R1040
                   LIB3170-076-Q1-K1-G3
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4580461
BLAST score
                   317
                   4.0e-29
E value
                   112
Match length
% identity
                   (AC006081) unknown protein [Arabidopsis thaliana]
NCBI Description
                   8054
Seq. No.
                   11378 2.R1040
Contig ID
                   hrw701060842.hl
5'-most EST
                   BLASTX
Method
                   g4580461
NCBI GI
BLAST score
                   147
                   2.0e-09
E value
```

Match length 77 32 % identity

(AC006081) unknown protein [Arabidopsis thaliana] NCBI Description

8055 Seq. No.

11386 1.R1040 Contig ID

5'-most EST LIB3049-028-Q1-E1-D1



```
BLASTX
Method
                  g2558660
NCBI GI
BLAST score
                   196
                   1.0e-14
E value
Match length
                   178
% identity
                   31
NCBI Description (AC002354) No definition line found [Arabidopsis thaliana]
                   8056
Seq. No.
                   11386 2.R1040
Contig ID
5'-most EST
                   eep700864234.h1
                   8057
Seq. No.
                   11389 1.R1040
Contig ID
                   LIB3167-078-P1-K2-C12
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1362103
BLAST score
                   816
                   3.0e-87
E value
                   192
Match length
% identity
                   81
                   ubiquitin conjugating enzyme - tomato
NCBI Description
                   >gi 886679 emb CAA58111 (X82938) ubiquitin conjugating
                   enzyme [Lycopersicon esculentum]
                   8058
Seq. No.
Contig ID
                   11389 2.R1040
                   pcp700990943.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1362103
BLAST score
                   450
E value
                   6.0e-45
Match length
                   103
% identity
                   82
                   ubiquitin conjugating enzyme - tomato
NCBI Description
                   >gi 886679 emb CAA58111 (X82938) ubiquitin conjugating
                   enzyme [Lycopersicon esculentum]
                   8059
Seq. No.
                   11396 1.R1040
Contig ID
                   LIB3040-052-Q1-E1-C5
5'-most EST
                   8060
Seq. No.
                   11400 1.R1040
Contig ID
5'-most EST
                   LIB3170-004-Q1-K1-D11
                   8061
Seq. No.
                   11407 1.R1040
Contig ID
5'-most EST
                   leu70\overline{1}154337.h1
Method
                   BLASTX
                   g3885941
NCBI GI
```

Method BLASTX
NCBI GI g3885941
BLAST score 312
E value 1.0e-28
Match length 85
% identity 69

NCBI Description (AF095454) PII protein [Ricinus communis]

Contig ID 11416_1.R1040

5'-most EST LIB3040-051-Q1-E1-G10

Seq. No. 8063

Contig ID 11419_1.R1040

5'-most EST LIB3040-051-Q1-E1-G2

Method BLASTX
NCBI GI g2160694
BLAST score 620
E value 9.0e-65
Match length 141
% identity 83

NCBI Description (U73528) B' regulatory subunit of PP2A [Arabidopsis

thaliana]

Seq. No. 8064

Contig ID 11427_1.R1040

5'-most EST LIB3040-044-Q1-E1-H3

Method BLASTX
NCBI GI g2398533
BLAST score 519
E value 2.0e-52
Match length 113
% identity 92

NCBI Description (Y13725) Transcription factor [Arabidopsis thaliana]

Seq. No. 8065

Contig ID 11427_2.R1040

5'-most EST uC-gmrominsoy258e01b1

Seq. No. 8066

Contig ID 11431_1.R1040 5'-most EST fde700875381.h1

Method BLASTX
NCBI GI g2894611
BLAST score 210
E value 7.0e-30
Match length 119
% identity 66

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8067

Contig ID 11433_1.R1040

5'-most EST LIB3107-063-Q1-K1-F1

Method BLASTX
NCBI GI g3695408
BLAST score 268
E value 3.0e-23
Match length 86
% identity 63

NCBI Description (AF096373) contains similarity to Solanum lycopersicum

(tomato) wound-induced protein (GB:X59882) [Arabidopsis thaliana] >gi_4538956_emb_CAB39780.1_ (AL049488) probable

wound-induced protein [Arabidopsis thaliana]

Seq. No. 8068



Contig ID 11433_2.R1040

5'-most EST LIB3040-051-Q1-E1-C2

Method BLASTX
NCBI GI g3695408
BLAST score 270
E value 1.0e-23
Match length 86
% identity 63

NCBI Description (AF096373) contains similarity to Solanum lycopersicum (tomato) wound-induced protein (GB:X59882) [Arabidopsis thaliana] >gi_4538956_emb_CAB39780.1_ (AL049488) probable

wound-induced protein [Arabidopsis thaliana]

Seq. No. 8069

Contig ID 11433 3.R1040

5'-most EST LIB3093-036-Q1-K1-A1

Seq. No. 8070

Contig ID 11437 1.R1040

5'-most EST LIB3040-051-Q1-E1-D10

Seq. No. 8071

Contig ID 11440_1.R1040 5'-most EST leu701156215.h1

Method BLASTN
NCBI GI g169364
BLAST score 111
E value 2.0e-55
Match length 210
% identity 89

NCBI Description Phaseolus vulgaris 5.8 Kb basic protein (PR4) gene, complete cds. >gi_217988_dbj_D12914_PHVPVPR4 Phaseolus

vulgaris PvPR4 mRNA for 5.8 kb basic protein, complete cds

Seq. No. 8072

Contig ID 11440_2.R1040

5'-most EST LIB3170-002-Q1-K1-H3

Method BLASTN
NCBI GI g169364
BLAST score 111
E value 2.0e-55
Match length 210
% identity 89

NCBI Description Phaseolus vulgaris 5.8 Kb basic protein (PR4) gene, complete cds. >gi 217988 dbj D12914_PHVPVPR4 Phaseolus

vulgaris PvPR4 mRNA for 5.8 kb basic protein, complete cds

Seq. No. 8073

Contig ID 11440 3.R1040

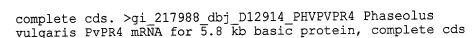
5'-most EST LIB3049-047-Q1-E1-A9

Method BLASTN
NCBI GI g169364
BLAST score 123
E value 1.0e-62
Match length 209
% identity 91

NCBI Description Phaseolus vulgaris 5.8 Kb basic protein (PR4) gene,

NCBI GI

BLAST score



```
8074
Seq. No.
Contig ID
                  11448 1.R1040
5'-most EST
                  LIB3049-042-Q1-E1-G9
                  BLASTX
Method
                  q3273828
NCBI GI
                  1597
BLAST score
                  0.0e + 00
E value
Match length
                  411
% identity
                  (AF068686) nodule-enhanced malate dehydrogenase [Glycine
NCBI Description
                  8075
Seq. No.
Contig ID
                  11448 2.R1040
5'-most EST
                  qsv701053423.h1
                  BLASTN
Method
                  g3273826
NCBI GI
                  147
BLAST score
                   6.0e-77
E value
                   318
Match length
                   91
% identity
                  Glycine max nodule-enhanced malate dehydrogenase (Mdh)
NCBI Description
                   gene, complete cds
                   8076
Seq. No.
                   11448 3.R1040
Contig ID
5'-most EST
                   qsv701043956.hl
Method
                   BLASTN
NCBI GI
                   q3273826
BLAST score
                   186
                   1.0e-100
E value
                   328
Match length
                   98
% identity
                  Glycine max nodule-enhanced malate dehydrogenase (Mdh)
NCBI Description
                   gene, complete cds
                   8077
Seq. No.
Contig ID
                   11468 1.R1040
                   LIB3040-047-Q1-E1-H6
5'-most EST
Method
                   BLASTX
                   q1173624
NCBI GI
                   308
BLAST score
                   5.0e-28
E value
                   71
Match length
                   77
% identity
                   (U34744) cytochrome P-450 [Phalaenopsis sp. 'hybrid
NCBI Description
                   SM9108']
                   8078
Seq. No.
                   11468 2.R1040
Contig ID
                   LIB3040-026-Q1-E1-H6
5'-most EST
                   BLASTX
Method
```

g2935525

148



```
2.0e-09
E value
Match length
                   76
% identity
                   49
NCBI Description (AF049067) cytochrome P450 [Pinus radiata]
                   8079
Seq. No.
                   11479 1.R1040
Contig ID
5'-most EST
                   LIB3040-050-Q1-E1-G7
Method
                   BLASTX
                   g2244935
NCBI GI
                   285
BLAST score
                   4.0e-25
E value
                   106
Match length
% identity
                   57
NCBI Description (Z97339) hypothetical protein [Arabidopsis thaliana]
                   8080
Seq. No.
                   11479 2.R1040
Contig ID
5'-most EST
                   LIB3170-016-Q1-K1-H1
                   BLASTX
Method
NCBI GI
                   g2244935
BLAST score
                   215
                   3.0e-17
E value
Match length
                   66
% identity
                   64
                  (Z97339) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   8081
Seq. No.
                   11481 1.R1040
Contig ID
5'-most EST
                   vzy700752646.h1
Method
                   BLASTX
                   q2494905
NCBI GI
BLAST score
                   174
E value
                   2.0e-12
Match length
                   90
                   46
% identity
NCBI Description
                   MICROTUBULE-ASSOCIATED PROTEIN YTM1 >gi_2132107_pir__S67174
                   hypothetical protein YOR272w - yeast (Saccharomyces
                   cerevisiae) >gi_1279699_emb_CAA61778_ (X89633) hypothetical
                   protein [Saccharomyces cerevisiae]
                   >gi_1420610_emb_CAA99497_ (Z75180) ORF YOR272w [Saccharomyces cerevisiae] >gi_1928989 (U92821)
                   microtubule-associated protein [Saccharomyces cerevisiae]
Seq. No.
                   8082
                   11483 1.R1040
Contig ID
5'-most EST
                   leu701149068.h1
Method
                   BLASTX
NCBI GI
                   q2289907
BLAST score
                   557
E value
                   6.0e-57
Match length
                   206
% identity
                   55
NCBI Description (AC002397) C2F [Mus musculus]
```

Contig ID 11483_2.R1040

8083

Seq. No.



```
leu701155519.h1
5'-most EST
Seq. No.
                  8084
Contig ID
                  11497 1.R1040
5'-most EST
                  LIB3073-008-Q1-K1-D4
                  BLASTX
Method
                  g2275216
NCBI GI
                  235
BLAST score
                  2.0e-19
E value
Match length
                  64
% identity
                  69
                  (AC002337) cytochrome c oxidase Vc subunit isolog
NCBI Description
                  [Arabidopsis thaliana]
Seq. No.
                  8085
Contig ID
                  11497 2.R1040
5'-most EST
                  LIB3170-054-Q1-K1-G12
Method
                  BLASTX
                  q2275216
NCBI GI
BLAST score
                  231
E value
                  3.0e-19
Match length
                  64
% identity
                  67
                  (AC002337) cytochrome c oxidase Vc subunit isolog
NCBI Description
                  [Arabidopsis thaliana]
                  8086
Seq. No.
Contig ID
                  11501 1.R1040
5'-most EST
                  LIB3170-015-Q1-K1-E8
                  8087
Seq. No.
                  11505 1.R1040
Contig ID
                  LIB3170-015-Q1-K1-E7
5'-most EST
                  8088
Seq. No.
Contig ID
                  11513 1.R1040
5'-most EST
                  LIB3040-061-Q1-E11-G7
                  BLASTX
Method
                  g3282674
NCBI GI
                  414
BLAST score
                  2.0e-40
E value
                  99
Match length
                  77
% identity
                  (AF036684) CCAAT-box binding factor HAP3 homolog
NCBI Description
                  [Arabidopsis thaliana]
                  8089
Seq. No.
Contig ID
                  11513 2.R1040
5'-most EST
                  LIB3040-026-Q1-E1-C6
Method
                  BLASTX
NCBI GI
                  g3282674
BLAST score
                  322
                  1.0e-29
E value
Match length
                  72
                  85
% identity
                  (AF036684) CCAAT-box binding factor HAP3 homolog
NCBI Description
```

1477

[Arabidopsis thaliana]



```
8090
Seq. No.
                  11517 1.R1040
Contig ID
                   epx701110032.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                   q3004560
                   668
BLAST score
                   5.0e-70
E value
                   149
Match length
                   85
% identity
                   (AC003673) putative ATP binding protein [Arabidopsis
NCBI Description
                   thaliana]
                   8091
Seq. No.
                   11517 2.R1040
Contig ID
5'-most EST
                   LIB3040-050-Q1-E1-C7
Method
                   BLASTX
                   q3004560
NCBI GI
```

287 BLAST score

7.0e-26 E value 75 Match length 75 % identity

(AC003673) putative ATP binding protein [Arabidopsis NCBI Description

thaliana]

8092 Seq. No.

11518 1.R1040 Contig ID

LIB3170-015-Q1-K1-E1 5'-most EST

BLASTX Method NCBI GI q3980254 BLAST score 299 2.0e-29 E value Match length 109 % identity

(AJ006053) peroxisomal membrane protein [Arabidopsis NCBI Description

thaliana]

8093 Seq. No.

11523 1.R1040 Contig ID $jex70\overline{0}903169.h1$ 5'-most EST

8094 Seq. No.

Contig ID 11525 1.R1040

LIB3040-049-Q1-E1-H12 5'-most EST

BLASTX Method g3024583 NCBI GI BLAST score 270 9.0e-24 E value 96 Match length 56 % identity

RT14 PROTEIN HOMOLOG >gi_2347196 (AC002338) RT14 protein NCBI Description

isolog [Arabidopsis thaliana]

8095 Seq. No.

11532 1.R1040 Contig ID hyd700729012.h1 5'-most EST

BLASTX Method



NCBI GI g1657617 BLAST score 604 E value 1.0e-62 Match length 135 % identity 83

NCBI Description (U72503) G2p [Arabidopsis thaliana] >gi_3068707 (AF049236)

putative nuclear DNA-binding protein G2p [Arabidopsis

thaliana]

Seq. No. 8096

Contig ID 11532 2.R1040

5'-most EST LIB3170-013-Q1-J1-D8

Method BLASTX
NCBI GI g1657617
BLAST score 416
E value 7.0e-41
Match length 105
% identity 78

NCBI Description (U72503) G2p [Arabidopsis thaliana] >gi_3068707 (AF049236)

putative nuclear DNA-binding protein G2p [Arabidopsis

thalianal

Seq. No. 8097

Contig ID 11533_1.R1040

5'-most EST LIB3040-050-Q1-E1-A5

Seq. No. 8098

Contig ID 11538 1.R1040

5'-most EST LIB3170-013-Q1-K1-D12

Seq. No. 8099

Contig ID 11538 2.R1040

5'-most EST LIB3109-050-Q1-K1-A1

Method BLASTX
NCBI GI g1706103
BLAST score 331
E value 2.0e-30
Match length 243
% identity 34

NCBI Description CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR, 100 KD

SUBUNIT (CPSF 100 KD SUBUNIT) >gi_1363022_pir__A56351 cleavage and polyadenylation specificity factor 100K chain - bovine >gi_599683_emb_CAA53535_ (X75931) Cleavage and Polyadenylation specificity factor (CPSF) 100kD subunit

[Bos taurus]

Seq. No. 8100

Contig ID 11557 1.R1040 5'-most EST kll701207775.h1

Method BLASTX
NCBI GI g3747132
BLAST score 489
E value 2.0e-49
Match length 124
% identity 71

NCBI Description (AF096776) expansin [Lycopersicon esculentum]



```
8101
Seq. No.
                   11561 1.R1040
Contig ID
                   vzy70\overline{0}755696.h1
5'-most EST
Method
                   BLASTX
                   g2244939
NCBI GI
BLAST score
                   203
                   1.0e-24
E value
Match length
                   181
% identity
                  (Z97339) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   8102
Seq. No.
                   11562 1.R1040
Contig ID
                   uaw70\overline{0}662405.h1
5'-most EST
Seq. No.
                   8103
                   11562 3.R1040
Contig ID
                   zpv700759748.h1
5'-most EST
                   8104
Seq. No.
Contig ID
                   11569 1.R1040
                   LIB3074-039-Q1-K1-F6
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4467116
BLAST score
                   312
E value
                   3.0e-28
Match length
                   134
                   46
% identity
                   (AL035538) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   8105
Seq. No.
                   11570 1.R1040
Contig ID
                   LIB3040-049-Q1-E1-E2
5'-most EST
                   8106
Seq. No.
                   11571 1.R1040
Contig ID
                   uC-gmrominsoy063e09b1
5'-most EST
                   BLASTX
Method
                   g3434975
NCBI GI
                   392
BLAST score
                   7.0e-38
E value
                   172
Match length
                   54
% identity
                   (AB008107) ethylene responsive element binding factor 5
NCBI Description
                    [Arabidopsis thaliana]
                   8107
Seq. No.
                   11571 2.R1040
Contig ID
5'-most EST
                   LIB3040-049-Q1-E1-E3
Method
                   BLASTX
NCBI GI
                   q3434975
                   210
BLAST score
                   1.0e-16
E value
Match length
                   141
```

[Arabidopsis thaliana]

44

% identity

NCBI Description

(AB008107) ethylene responsive element binding factor 5



8108 Seq. No. 11577 1.R1040 Contig ID fC-gmse700669227a3 5'-most EST BLASTX Method q3128167 NCBI GI 796 BLAST score 1.0e-84 E value 483 Match length 43 % identity (AC004521) hypothetical protein [Arabidopsis thaliana] NCBI Description >gi 3212874 (AC004005) hypothetical protein [Arabidopsis thaliana] 8109 Seq. No. 11577 2.R1040 Contig ID 5'-most EST LIB3170-015-Q1-K1-D6 8110 Seq. No. 11589 1.R1040 Contig ID gsv701054713.hl 5'-most EST BLASTX Method q3421109 NCBI GI 992 BLAST score 1.0e-108 E value 204 Match length % identity (AF043533) 20S proteasome beta subunit PBC2 [Arabidopsis NCBI Description thaliana] 8111 Seq. No. 11589 2.R1040 Contig ID LIB3106-114-Q1-K1-G7 5'-most EST BLASTX Method q3421109 NCBI GI 776 BLAST score 1.0e-82 E value Match length 171 % identity (AF043533) 20S proteasome beta subunit PBC2 [Arabidopsis NCBI Description thaliana] Seq. No. 8112 11589 3.R1040 Contig ID $LIB30\overline{4}9-056-Q1-E1-A10$ 5'-most EST BLASTX Method g3421109 NCBI GI BLAST score 340 3.0e-42E value Match length 98 % identity (AF043533) 20S proteasome beta subunit PBC2 [Arabidopsis NCBI Description

Seq. No. 8113 Contig ID 1158

Contig ID 11589_4.R1040 5'-most EST k11701204594.h2

thaliana]

```
Method
                  BLASTX
NCBI GI
                  g3421109
BLAST score
                  616
                  3.0e-64
E value
                  127
Match length
                  93
% identity
                  (AF043533) 20S proteasome beta subunit PBC2 [Arabidopsis
NCBI Description
                  thalianal
Seq. No.
                  8114
Contig ID
                  11601 1.R1040
                  gsv701046451.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3850571
BLAST score
                   427
E value
                   8.0e-42
Match length
                   92
% identity
                   88
                   (AC005278) Similar to gb U85207 snRNP core Sm protein
NCBI Description
                  homolog Sm-X5 from Mus musculus. EST gb AA612141 comes
                   from this gene. [Arabidopsis thaliana]
Seq. No.
                  8115
                  11601 2.R1040
Contig ID
5'-most EST
                  LIB3050-004-Q1-E1-E9
```

BLASTX Method g3850571 NCBI GI BLAST score 233 E value 3.0e-29

Match length 79 87 % identity NCBI Description

(ACO05278) Similar to gb U85207 snRNP core Sm protein homolog Sm-X5 from Mus musculus. EST gb AA612141 comes from this gene. [Arabidopsis thaliana]

Seq. No. 8116

11603 1.R1040 Contig ID

5'-most EST jC-gmro02910060d01a1

Seq. No. 8117

11613_1.R1040 Contig ID

5'-most EST LIB3074-004-Q1-K1-E9

8118 Seq. No.

11620 1.R1040 Contig ID

5'-most EST LIB3170-016-Q1-J1-D4

Seq. No. 8119

Contig ID 11629 1.R1040

5'-most EST LIB3049-005-Q1-E1-G9

Method BLASTX NCBI GI g3044218 BLAST score 581 3.0e-60 E value Match length 130 % identity 82

NCBI Description (AF057144) signal peptidase [Arabidopsis thaliana]

Contig ID

```
8120
Seq. No.
                  11631 1.R1040
Contig ID
5'-most EST
                  LIB3087-008-Q1-K1-G5
                  8121
Seq. No.
                  11632 1.R1040
Contig ID
5'-most EST
                  LIB3040-045-Q1-E1-C11
Seq. No.
                  8122
                  11635 1.R1040
Contig ID
5'-most EST
                  uC-gmropic036b09b1
Method
                  BLASTX
NCBI GI
                  q4091080
BLAST score
                  528
E value
                  1.0e-53
Match length
                  201
                  52
% identity
NCBI Description (AF045571) nucleic acid binding protein [Oryza sativa]
                  8123
Seq. No.
Contia ID
                  11635 2.R1040
5'-most EST
                  LIB3049-033-Q1-E1-D11
Method
                  BLASTX
NCBI GI
                  g166410
BLAST score
                  319
E value
                  2.0e-29
Match length
                  91
% identity
                  71
NCBI Description
                  (L07291) Alfin-1 [Medicago sativa]
                  8124
Seq. No.
Contig ID
                  11635 3.R1040
5'-most EST
                  leu701147577.h1
Method
                  BLASTX
                  g166410
NCBI GI
BLAST score
                  287
E value
                  9.0e-26
                  62
Match length
                  85
% identity
NCBI Description
                 (L07291) Alfin-1 [Medicago sativa]
Seq. No.
                  8125
Contig ID
                  11637 1.R1040
                  LIB3106-061-Q1-K1-B6
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2388578
BLAST score
                  861
E value
                  1.0e-92
Match length
                  202
% identity
                  81
NCBI Description
                  (AC000098) Similar to Mycobacterium RlpF (gb Z84395). ESTs
                   gb_T75785,gb_R30580,gb_T04698 come from this gene.
                   [Arabidopsis thaliana]
Seq. No.
                  8126
```

11640 1.R1040

```
qsv701047317.hl
5'-most EST
                   8127
Seq. No.
                   11641 1.R1040
Contiq ID
                   jC-gmf102220113h02a1
5'-most EST
                   BLASTX
Method
                   q4512708
NCBI GI
BLAST score
                   348
                   9.0e-33
E value
Match length
                   84
                   76
% identity
NCBI Description (AC006569) putative glutaredoxin [Arabidopsis thaliana]
                   8128
Seq. No.
                   11641 2.R1040
Contig ID
                   LIB30\overline{4}0-047-Q1-E1-H9
5'-most EST
                   BLASTX
Method
                   g4512708
NCBI GI
                   326
BLAST score
                   3.0e - 30
E value
                   83
Match length
                   75
% identity
NCBI Description (AC006569) putative glutaredoxin [Arabidopsis thaliana]
                   8129
Seq. No.
                   11656 1.R1040
Contig ID
                   LIB3040-047-Q1-E1-F6
5'-most EST
                   BLASTX
Method
                   g4455325
NCBI GI
                   415
BLAST score
                   2.0e-40
E value
                   142
Match length
% identity
                  (AL035525) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    11660 1.R1040
Contig ID
 5'-most EST
                   LIB3051-031-Q1-K1-D8
 Seq. No.
                    11660 3.R1040
 Contig ID
                    LIB3170-014-Q1-K1-C4
 5'-most EST
 Seq. No.
                    11661 1.R1040
 Contig ID
                    LIB3040-047-Q1-E1-G12
 5'-most EST
                    BLASTX
 Method
                    g4490704
 NCBI GI
 BLAST score
                    548
                    6.0e-56
 E value
                    215
 Match length
                    52
 % identity
                   (AL035680) putative protein [Arabidopsis thaliana]
 NCBI Description
                    8133
 Seq. No.
                    11668 1.R1040
 Contig ID
```

1484

LIB3139-059-P1-N1-G6

5'-most EST



```
BLASTX
Method
                  g3080401
NCBI GI
                  920
BLAST score
                  1.0e-99
E value
                  233
Match length
                   72
% identity
                   (AL022603) putative protein [Arabidopsis thaliana]
NCBI Description
                  >gi_4455265_emb_CAB36801.1_ (AL035527) putative protein
                   [Arabidopsis thaliana]
                   8134
Seq. No.
                   11672 1.R1040
Contig ID
                  LIB3094-047-Q1-K1-A4
5'-most EST
Method
                   BLASTX
                   g2865175
NCBI GI
                   556
BLAST score
                   8.0e-57
E value
                   154
Match length
% identity
                   68
                  (AB010945) AtRerlA [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   8135
                   11672 2.R1040
Contig ID
                   vzy70\overline{0}755937.h1
5'-most EST
                   BLASTX
Method
                   g2865177
NCBI GI
                   555
BLAST score
                   8.0e-57
E value
                   183
Match length
                   57
% identity
                   (AB010946) AtRer1B [Arabidopsis thaliana]
NCBI Description
                   8136
Seq. No.
                   11672 3.R1040
Contig ID
                   jC-gmst02400015f10a1
5'-most EST
                   BLASTX
Method
                   q2865175
NCBI GI
BLAST score
                   168
                   2.0e-11
E value
                   49
Match length
% identity
                   (AB010945) AtRer1A [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    8137
                    11672 4.R1040
Contig ID
                   LIB3094-047-Q1-K1-A3
 5'-most EST
                   BLASTX
Method
                    g2865175
NCBI GI
                    390
BLAST score
                    1.0e-37
E value
                    119
Match length
 % identity
                    61
                    (AB010945) AtRer1A [Arabidopsis thaliana]
 NCBI Description
```

Seq. No. 8138

Contig ID 11672_5.R1040

5'-most EST LIB3049-001-Q1-E1-H8

1485

```
Method BLASTX

NCBI GI g2865175

BLAST score 197

E value 3.0e-15

Match length 83

% identity 48

NCBI Description (AB010945) AtRerlA [Arabidopsis thaliana]
```

Seq. No.
Contig ID
5'-most EST

8139 11672_6.R1040 rlr700897411.h1

Seq. No.

8140

Contig ID 11673_1.R1040

5'-most EST

LIB3170-015-Q1-K1-C12

Seq. No.

8141

Contig ID

11680_1.R1040

5'-most EST Method LIB3072-023-Q1-E1-H11

Method BLASTX
NCBI GI g1723738
BLAST score 222
E value 5.0e-18
Match length 103
% identity 30

NCBI Description HYPOTHETI

HYPOTHETICAL 55.1 KD PROTEIN IN TRX1-RTA1 INTERGENIC REGION >gi_2131699_pir__S64534 hypothetical protein YGR211w - yeast (Saccharomyces cerevisiae) >gi_1323379_emb_CAA97238_(Z72996) ORF YGR211w [Saccharomyces cerevisiae] >gi_3510466 (AF019769) zinc finger protein [Saccharomyces cerevisiae]

Seq. No. 8142

Contig ID 11693 1.R1040

5'-most EST LIB3040-047-Q1-E1-C4

Seq. No. 8143

Contig ID 11693 2.R1040

5'-most EST LIB3093-037-Q1-K1-G5

Seq. No. 8144

Contig ID 11694 1.R1040 5'-most EST ncj700979372.h1 Method BLASTX

NCBI GI g4455207
BLAST score 443
E value 7.0e-44
Match length 91
% identity 93

NCBI Description (AL035440) ubiquitin-like protein [Arabidopsis thaliana]

Seq. No. 8145

Contig ID 11695_1.R1040

5'-most EST LIB $30\overline{7}2-057-Q1-K1-G1$

Method BLASTX
NCBI GI g1082054
BLAST score 153
E value 1.0e-09



```
142
Match length
                   30
% identity
NCBI Description (Z49859) copper transporter protein [Arabidopsis thaliana]
                   8146
Seq. No.
Contig ID
                   11695 2.R1040
                  q4284\overline{4}06
5'-most EST
                   8147
Seq. No.
                   11696 1.R1040
Contig ID
5'-most EST
                   LIB3040-047-Q1-E1-C7
                   8148
Seq. No.
                   11697 1.R1040
Contig ID
                   LIB3170-085-Q1-J1-C2
5'-most EST
Seq. No.
                   8149
Contig ID
                   11700 1.R1040
                   LIB3050-015-Q1-E1-C4
5'-most EST
Method
                   BLASTN
                   q3449334
NCBI GI
BLAST score
                   46
                   1.0e-16
E value
                   126
Match length
% identity
                   84
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MYH9, complete sequence [Arabidopsis thaliana]
                   8150
Seq. No.
Contig ID
                   11707 1.R1040
                   LIB3170-015-Q1-K1-C6
5'-most EST
                   8151
Seq. No.
                   11708 1.R1040
Contig ID
                   LIB3170-016-Q1-K1-C7
5'-most EST
                   8152
Seq. No.
                   11712 1.R1040
Contig ID
                   LIB3170-015-Q1-J1-C7
5'-most EST
                   BLASTX
Method
                   g4006878
NCBI GI
                   140
BLAST score
                   1.0e-08
E value
                   47
Match length
% identity
                   53
                   (Z99707) MAP3K-like protein kinase [Arabidopsis thaliana]
NCBI Description
                   8153
Seq. No.
                   11716 1.R1040
Contig ID
```

LIB3040-046-Q1-E1-G9 5'-most EST

Seq. No. 8154

11717 1.R1040 Contig ID

5'-most EST LIB3040-047-Q1-E1-A7

Seq. No.

11727_1.R1040 Contig ID



5'-most EST LIB3040-046-Q1-E1-F9

Seq. No. 8156

Contig ID 11732_1.R1040

5'-most EST LIB3170-016-Q1-K1-C2

Seq. No. 8157

Contig ID 11734 1.R1040

5'-most EST LIB3170-015-Q1-K1-C2

Seq. No. 8158

Contig ID 11734 5.R1040 5'-most EST leu701155165.h1

Seq. No. 8159

Contig ID 11734_8.R1040 5'-most EST gsv701044315.h1

Seq. No. 8160

Contig ID 11737_1.R1040

5'-most EST jC-gmle01810092d05a1

Method BLASTX
NCBI GI g1699023
BLAST score 215
E value 9.0e-17
Match length 225
% identity 36

NCBI Description (U78866) putative arginine-aspartate-rich RNA binding

protein [Arabidopsis thaliana] >gi_1699051 (U78867) putative aspartate-arginine-rich mRNA binding protein

[Arabidopsis thaliana]

Seq. No. 8161

Contig ID 11740_1.R1040 5'-most EST gsv701044686.h1

Seq. No. 8162

Contig ID 11740_2.R1040

5'-most EST LIB3170-013-Q1-K1-E2

Seq. No. 8163

Contig ID 11746_1.R1040

5'-most EST LIB3109-001-Q1-K4-D1

Method BLASTX
NCBI GI g2760320
BLAST score 545
E value 2.0e-55
Match length 223
% identity 48

NCBI Description (AC002130) F1N21.4 [Arabidopsis thaliana]

Seq. No. 8164

Contig ID 11755_1.R1040

5'-most EST LIB3040-029-Q1-E1-E4

Method BLASTN
NCBI GI g3204128
BLAST score 200

1488



E value 1.0e-108
Match length 404
% identity 87

NCBI Description Cicer arietinum mRNA for histone H2A

Seq. No. 8165

Contig ID 11758_1.R1040

5'-most EST LIB3040-046-Q1-E1-E5

Method BLASTX
NCBI GI g2281088
BLAST score 350
E value 6.0e-33
Match length 159
% identity 42

NCBI Description (AC002333) indole-3-acetate beta-glucosyltransferase isolog

[Arabidopsis thaliana]

Seq. No. 8166

Contig ID 11762_1.R1040

5'-most EST LIB3040-045-Q1-E1-H9

Method BLASTN
NCBI GI g633606
BLAST score 44
E value 3.0e-15
Match length 183
% identity 85

NCBI Description P.sativum mRNA for chloroplastic outer envelope membrane

protein (OEP75)

Seq. No. 8167

Contig ID 11762_2.R1040

5'-most EST LIB3093-012-Q1-K1-H10

Method BLASTN
NCBI GI g633606
BLAST score 57
E value 4.0e-23
Match length 230
% identity 86

NCBI Description P.sativum mRNA for chloroplastic outer envelope membrane

protein (OEP75)

Seq. No. 8168

Contig ID 11762_3.R1040 5'-most EST gsv701050073.h1

Seq. No. 8169

Contig ID 11762 5.R1040

5'-most EST jC-gmle01810088c07d1

Method BLASTN
NCBI GI g633606
BLAST score 34
E value 1.0e-09
Match length 50
% identity 92

NCBI Description P.sativum mRNA for chloroplastic outer envelope membrane

protein (OEP75)



Seq. No. 8170

Contig ID 11763_1.R1040 5'-most EST pmv700888903.h1

Method BLASTX
NCBI GI g3096931
BLAST score 301
E value 4.0e-27
Match length 78
% identity 71

NCBI Description (AL023094) putative ribosomal protein S16 [Arabidopsis

thaliana]

Seq. No. 8171

Contig ID 11768_1.R1040 5'-most EST zzp700834271.h1

Seq. No. 8172

Contig ID 11774_1.R1040

5'-most EST LIB3040-046-Q1-E1-B12

Seq. No. 8173

Contig ID 11785_1.R1040

5'-most EST LIB3040-045-Q1-E1-G8

Seq. No. 8174

Contig ID 11788_1.R1040

5'-most EST LIB3040-045-Q1-E1-H12

Method BLASTN
NCBI GI g2244788
BLAST score 56
E value 1.0e-22
Match length 197
% identity 88

NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I contig

fragment No

Seq. No. 8175

Contig ID 11788_2.R1040

5'-most EST LIB3040-016-Q1-E1-G12

Method BLASTN
NCBI GI g2244788
BLAST score 60
E value 4.0e-25
Match length 197
% identity 89

NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I contig

fragment No

Seq. No. 8176

Contig ID 11788_3.R1040 5'-most EST zsg701127979.h1

Method BLASTN
NCBI GI g2244788
BLAST score 47
E value 2.0e-17
Match length 133
% identity 89



NCBI Description Arabidopsis thaliana DNA chromosome 4, ESSA I contig fragment No

Seq. No. 8177

Contig ID 11796_1.R1040

5'-most EST LIB3170-031-Q1-K1-H3

Method BLASTN
NCBI GI g3241927
BLAST score 62
E value 8.0e-26
Match length 338
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTE17, complete sequence [Arabidopsis thaliana]

Seq. No. 8178

Contig ID 11796_2.R1040 5'-most EST kmv700740305.h1

Method BLASTN
NCBI GI g3241927
BLAST score 37
E value 4.0e-11
Match length 138
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTE17, complete sequence [Arabidopsis thaliana]

Seq. No. 8179

Contig ID 11796_3.R1040

5'-most EST LIB3093-003-Q1-K1-F4

Method BLASTN
NCBI GI g3241927
BLAST score 33
E value 1.0e-08
Match length 150
% identity 85

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MTE17, complete sequence [Arabidopsis thaliana]

Seq. No. 8180

Contig ID 11796_5.R1040

5'-most EST LIB30 $\overline{5}1-115-Q1-K1-G5$

Seq. No. 8181

Contig ID 11796_6.R1040 5'-most EST asn701142715.h1

Seq. No. 8182

Contig ID 11803 1.R1040

5'-most EST LIB3074-033-Q1-K1-H3

Method BLASTN
NCBI GI g343492
BLAST score 107
E value 3.0e-53
Match length 188
% identity 94

NCBI Description Tobacco Leu-tRNA and Phe-tRNA genes



8183 Seq. No.

11805 1.R1040 Contig ID

5'-most EST LIB3106-043-Q1-K1-G7

Seq. No. 8184

Contig ID 11807 1.R1040

LIB3040-045-Q1-E1-F6 5'-most EST

BLASTN Method NCBI GI g1370179 BLAST score 415 E value 0.0e+00Match length 591 93 % identity

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB5B

Seq. No.

8185 Contig ID 11807 2.R1040

5'-most EST

uC-gmronoir038c03b1

Method BLASTX NCBI GI q4559380 BLAST score 315 3.0e-29 E value Match length 89 62 % identity

NCBI Description (AC006526) putative auxin-responsive GH3 protein

[Arabidopsis thaliana]

8186 Seq. No.

Contig ID 11812 1.R1040

5'-most EST LIB3040-045-Q1-E1-C2

Method BLASTX NCBI GI q3249105 BLAST score 424 E value 3.0e-41 Match length 188 % identity 53

(AC003114) Contains similarity to protein phosphatase 2C NCBI Description (ABI1) gb X78886 from A. thaliana. [Arabidopsis thaliana]

Seq. No.

11812 2.R1040 Contig ID

5'-most EST LIB3109-017-Q1-K1-H8

8187

8188 Seq. No.

Contig ID 11818 1.R1040

5'-most EST LIB3106-097-01-K1-E8

Seq. No. 8189

11818 2.R1040 Contig ID 5'-most EST $k1170\overline{1}214013.h1$

8190 Seq. No.

11819 1.R1040 Contig ID

5'-most EST LIB3094-077-Q1-K1-D2

Seq. No. 8191



```
Contig ID
                  11821 1.R1040
5'-most EST
                  LIB3040-045-Q1-E1-D11
Method
                  BLASTN
NCBI GI
                  g470126
BLAST score
                  36
E value
                  1.0e-10
                  121
Match length
% identity
                  91
NCBI Description N.tabacum (cv.Samsun NN) L19 mRNA for ribosomal protein L19
Seq. No.
                  8192
Contig ID
                  11823 1.R1040
5'-most EST
                  LIB3170-068-Q1-J1-B9
Seq. No.
                  8193
Contig ID
                  11837 1.R1040
5'-most EST
                  kmv700742945.h1
Method
                  BLASTX
NCBI GI
                  q3915196
BLAST score
                  227
E value
                  2.0e-23
Match length
                  113
% identity
                  51
NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-C (UBIQUITIN-PROTEIN
                  LIGASE) (UBIQUITIN CARRIER PROTEIN) >gi 1493838 (U52949)
                  cyclin-specific ubiquitin carrier protein E2-C [Spisula
                  solidissima]
                  8194
Seq. No.
                  11847 1.R1040
Contig ID
5'-most EST
                  LIB3107-062-Q1-K1-G9
Method
                  BLASTN
NCBI GI
                  g11957
BLAST score
                  80
E value
                  9.0e-37
Match length
                  120
% identity
                  92
NCBI Description Rice complete chloroplast genome
Seq. No.
                  8195
Contig ID
                  11850 1.R1040
                  LIB3040-044-Q1-E1-H4
5'-most EST
Method
                  BLASTX
NCBI GI
                  q4115916
BLAST score
                  191
E value
                  2.0e-14
Match length
                  84
                  58
% identity
NCBI Description
                  (AF118222) F3H7.9 gene product [Arabidopsis thaliana]
                  >gi_4539441_emb_CAB40029.1_ (AL049523) putative protein
                   [Arabidopsis thaliana]
```

Seq. No. 8196

Contig ID 11854_1.R1040

5'-most EST jC-gmro02910016f04d1

Seq. No. 8197



```
11858 1.R1040
Contig ID
5'-most EST
                  LIB3051-025-Q1-K1-D2
Seq. No.
                  8198
                  11861 1.R1040
Contig ID
                  trc700564152.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  q1173257
                  1195
BLAST score
                  1.0e-131
E value
Match length
                  264
% identity
                  86
                 40S RIBOSOMAL PROTEIN S4 >qi 1076674 pir S47642 ribosomal
NCBI Description
                  protein S4 - potato >gi 457803 emb CAA54095 (X76651)
                  ribosomal protein S4 [Solanum tuberosum]
Seq. No.
                  8199
Contig ID
                  11868 1.R1040
5'-most EST
                  leu701151162.h1
Method
                  BLASTX
NCBI GI
                  g2407800
                  570
BLAST score
                  9.0e-59
E value
                  128
Match length
                  90
% identity
                 (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
NCBI Description
                  8200
Seq. No.
                  11868 2.R1040
Contig ID
5'-most EST
                  zsq701117481.h1
Method
                  BLASTX
NCBI GI
                  q2407800
BLAST score
                  440
E value
                  9.0e-44
Match length
                  114
% identity
                  78
NCBI Description (Y12575) histone H2A.F/Z [Arabidopsis thaliana]
Seq. No.
                  8201
                  11868 3.R1040
Contig ID
5'-most EST
                  LIB3040-044-Q1-E1-G2
                  BLASTX
Method
NCBI GI
                  g2407800
BLAST score
                  310
```

2.0e-28 E value 97 Match length % identity 68

NCBI Description (Y12575) histone H2A.F/Z [Arabidopsis thaliana]

8202 Seq. No. Contig ID

11869 1.R1040

5'-most EST LIB3040-044-Q1-E1-G4

8203 Seq. No.

Contig ID 11871 1.R1040

LIB3040-044-Q1-E1-G7 5'-most EST

Method BLASTX

```
ŵ≠ = *
                           q2243118
        NCBI GI
        BLAST score
                           550
        E value
                           2.0e-56
                           207
        Match length
                           55
        % identity
        NCBI Description (Y10984) glutathione synthetase [Brassica juncea]
        Seq. No.
                           8204
                           11871 2.R1040
        Contig ID
        5'-most EST
                          LIB3109-029-Q1-K1-D1
        Method
                          BLASTX
        NCBI GI
                           g2243118
        BLAST score
                          133
        E value
                          2.0e-13
                          129
        Match length
        % identity
                           40
        NCBI Description (Y10984) glutathione synthetase [Brassica juncea]
                           8205
        Seq. No.
        Contig ID
                           11875 1.R1040
        5'-most EST
                          LIB3040-044-Q1-E1-D12
        Seq. No.
                           8206
        Contig ID
                           11885 1.R1040
        5'-most EST
                           jC-gmle01810086e10d1
        Method
                           BLASTX
                           g2924258
        NCBI GI
        BLAST score
                           2995
                           0.0e + 00
        E value
        Match length
                           751
        % identity
                           79
        NCBI Description (Z00044) RNA polymerase beta' subunit [Nicotiana tabacum]
        Seq. No.
                           8207
                           11895 1.R1040
        Contig ID
        5'-most EST
                           LIB3106-105-Q1-K1-H3
        Method
                          BLASTX
        NCBI GI
                           q2342735
        BLAST score
                           480
                           5.0e-48
        E value
        Match length
                           124
                           80
        % identity
        NCBI Description (AC002341) unknown protein [Arabidopsis thaliana]
                           8208
        Seq. No.
                           11900 1.R1040
        Contig ID
        5'-most EST
                           dpv701099491.h1
                           BLASTX
        Method
                           q4566614
        NCBI GI
        BLAST score
                           564
        E value
                           1.0e-57
```

Match length 130

% identity

NCBI Description (AF112887) actin depolymerizing factor [Populus alba x

Populus tremula]

Seq. No. 8209



Contig ID 11910_1.R1040

5'-most EST LIB3107-010-Q1-K1-E6

Method BLASTX
NCBI GI g400983
BLAST score 640
E value 1.0e-66
Match length 215
% identity 63

% identity 63 NCBI Description 50S RIBOSOMAL PROTEIN L11, CHLOROPLAST PRECURSOR (CL11)

>gi_279648_pir__R5SP11 ribosomal protein L11 precursor spinach >gi_21313_emb_CAA39950_ (X56615) ribosomal protein

L11 [Spinacia oleracea]

Seq. No. 8210

Contig ID 11925_1.R1040 5'-most EST vzy700756050.h1

Seq. No. 8211

Contig ID 11926_1.R1040

5'-most EST LIB3040-043-Q1-E1-F9

Seq. No. 8212

Contig ID 11928_1.R1040 5'-most EST gsv701052309.h1

Seq. No. 8213

Contig ID 11930_1.R1040

5'-most EST LIB3040-043-Q1-E1-G2

Method BLASTX
NCBI GI g2088651
BLAST score 237
E value 2.0e-26
Match length 155
% identity 48

NCBI Description (AF002109) hypersensitivity-related gene 201 isolog

[Arabidopsis thaliana]

Seq. No. 8214

Contig ID 11933_1.R1040

5'-most EST LIB3040-017-Q1-E1-H7

Method BLASTX
NCBI GI g3874563
BLAST score 247
E value 5.0e-21
Match length 87
% identity 52

NCBI Description (Z81042) similar to Yeast hypothetical protein YEY6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk206h5.5

comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA ... >gi 3924825 emb CAB05549 (Z83113) similar to Yeast hypothetical protein YEY6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk206h5.5 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5

comes from this gene; cDNA

Seq. No. 8215

Match length

% identity

131

42

```
11937 1.R1040
Contig ID
                  taw700654539.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4337206
BLAST score
                  345
                  2.0e-32
E value
Match length
                  111
% identity
                   (AC006403) putative replication factor-A protein
NCBI Description
                   [Arabidopsis thaliana]
                  8216
Seq. No.
                   11938 1.R1040
Contig ID
                  jC-gmst02400062f11d1
5'-most EST
Seq. No.
                   8217
                   11940 1.R1040
Contig ID
                   gsv701045940.hl
5'-most EST
Method
                  BLASTX
                   g1173074
NCBI GI
BLAST score
                   179
                   9.0e-13
E value
Match length
                   111
% identity
                   37
                   60S ACIDIC RIBOSOMAL PROTEIN P2 (ALLERGEN CLA H 3) (CLA H
NCBI Description
                   III) >gi_1086177_pir__S41866 ribosomal P2 protein -
                   Cladosporium herberum >gi_452606_emb_CAA54470_ (X77253)
                   ribosomal P2 protein [Cladosporium herbarum]
                   8218
Seq. No.
                   11952 1.R1040
Contig ID
                   LIB3072-012-Q1-E1-C7
5'-most EST
                   8219
Seq. No.
                   11955_1.R1040
Contig ID
                   sat701011736.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2398521
BLAST score
                   218
                   3.0e-22
E value
Match length
                   219
% identity
                   42
                  (Y13720) transcription factor [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   8220
                   11955 2.R1040
Contig ID
                   LIB3040-043-Q1-E1-B11
5'-most EST
                   8221
Seq. No.
                   11959 1.R1040
Contig ID
5'-most EST
                   rlr700900885.h1
                   BLASTX
Method
NCBI GI
                   q4115735
BLAST score
                   257
                   5.0e-22
E value
```

```
NCBI Description (AB017627) acetyltransferase [Candida albicans]
                  8222
Seq. No.
Contig ID
                  11960 1.R1040
5'-most EST
                  LIB3040-043-Q1-E1-C10
Method
                  BLASTX
                  q728827
NCBI GI
BLAST score
                  202
                  9.0e-16
E value
Match length
                  90
% identity
                  50
                  ALLANTOINASE PRECURSOR >gi 458126 (U03471) allantoinase
NCBI Description
                  [Rana catesbeiana]
                  8223
Seq. No.
Contig ID
                  11963 1.R1040
5'-most EST
                  LIB3040-043-Q1-E1-C4
Method
                  BLASTX
                  q1001630
NCBI GI
BLAST score
                  330
                   2.0e-30
E value
Match length
                   95
% identity
                   62
                  (D64002) hypothetical protein [Synechocystis sp.]
NCBI Description
Seq. No.
                   8224
                   11968 1.R1040
Contig ID
5'-most EST
                   fua701039084.hl
Method
                   BLASTX
                   q3367591
NCBI GI
                   434
BLAST score
                   6.0e-43
E value
                   104
Match length
                   83
% identity
NCBI Description (AL031135) putative protein [Arabidopsis thaliana]
                   8225
Seq. No.
                   11969 1.R1040
Contig ID
                   pmv700894162.hl
5'-most EST
                   BLASTX
Method
                   g3292849
NCBI GI
                   549
BLAST score
                   3.0e-56
E value
                   141
Match length
% identity
                  (AJ007582) arginine methyltransferase [Arabidopsis
NCBI Description
                   thaliana]
                   8226
Seq. No.
Contig ID
                   11971 1.R1040
5'-most EST
                   LIB3040-001-Q1-E1-H10
Method
                   BLASTX
NCBI GI
                   q730922
BLAST score
                   1445
```

0.0e + 00

517

64

E value Match length

% identity



T-COMPLEX PROTEIN 1, ZETA SUBUNIT (TCP-1-ZETA) (CCT-ZETA) NCBI Description (TCP20) (HTR3) >gi 627402_pir__S48087 chaperonin TCP20 human >gi 517065 (L27706) chaperonin-like protein [Homo sapiens] >gi_4502643_ref_NP_001753.1_pCCT6_ chaperonin containing T-complex subunit 8227 Seq. No. 11975 1.R1040 Contig ID zhf700957834.h1 5'-most EST 8228 Seq. No. 11978 1.R1040 Contig ID LIB3139-009-P1-N1-C11 5'-most EST 8229 Seq. No. Contig ID 11983 1.R1040 LIB3170-015-Q1-K1-A9 5'-most EST 8230 Seq. No. 11985 1.R1040 Contig ID LIB3040-043-Q1-E1-A7 5'-most EST 8231 Seq. No. 11992 1.R1040 Contig ID zpv700759655.h1 5'-most EST BLASTX Method NCBI GI q3513744 332 BLAST score 7.0e-31 E value Match length 143 33 % identity (AF080118) contains similarity to Medicago truncatula MtN3 NCBI Description (GB:Y08726) [Arabidopsis thaliana] 8232 Seq. No. 11993 1.R1040 Contig ID LIB3040-042-Q1-E1-F12 5'-most EST 8233 Seq. No. 11998 1.R1040 Contig ID 5'-most EST LIB3040-042-Q1-E1-F9 BLASTX Method q4102839 NCBI GI 303 BLAST score 1.0e-27 E value 88 Match length % identity 64 (AF016713) LeOPT1 [Lycopersicon esculentum] NCBI Description

Seq. No. 8234

Contig ID 12008_1.R1040

5'-most EST LIB3040-042-Q1-E1-B6

Method BLASTX
NCBI GI g2736149
BLAST score 151
E value 1.0e-09
Match length 116

 Seq. No.
 8235

 Contig ID
 12011_1.R1040

 5'-most EST
 seb700648931.h1

 Method
 BLASTX

 NCBI GI
 g3335336

 RLAST score
 165

NCBI GI g3335330 BLAST score 165 E value 4.0e-11 Match length 63 % identity 49

NCBI Description (AC004512) Contains similarity to DnaJ gene YM8520.10 gb_825566 from from S. cerevisiae cosmid gb_Z49705. ESTs

gb_Z47720 and gb_Z29879 come from this gene. [Arabidopsis thaliana]

thaliana]

Seq. No. 8236

Contig ID 12018_1.R1040 5'-most EST zhf700954547.h1

Method BLASTX
NCBI GI g1172556
BLAST score 315
E value 7.0e-29
Match length 111
% identity 52

NCBI Description 36 KD OUTER MITOCHONDRIAL MEMBRANE PROTEIN PORIN

(VOLTAGE-DEPENDENT ANION-SELECTIVE CHANNEL PROTEIN) (VDAC) (POM 36) >gi_629729_pir__S46925 porin II, 36K - potato >gi_1076681_pir__B55364 porin (clone pPOM 36.2) - potato mitochondrion >gi_515360_emb_CAA56600_(X80387) 36kDA porin

II [Solanum tuberosum]

Seq. No. 8237

Contig ID 12018_2.R1040 5'-most EST gsv701052324.h1

Method BLASTX
NCBI GI g1172556
BLAST score 142
E value 4.0e-09
Match length 48
% identity 52

NCBI Description 36 KD OUTER MITOCHONDRIAL MEMBRANE PROTEIN PORIN

(VOLTAGE-DEPENDENT ANION-SELECTIVE CHANNEL PROTEIN) (VDAC) (POM 36) >gi_629729_pir__S46925 porin II, 36K - potato >gi_1076681_pir__B55364 porin (clone pPOM 36.2) - potato mitochondrion >gi_515360_emb_CAA56600_(X80387) 36kDA porin

II [Solanum tuberosum]

Seq. No. 8238

Contig ID 12028 1.R1040



5'-most EST LIB3040-041-Q1-E1-H9 Method BLASTX

NCBI GI q1076414 BLAST score 641 E value 5.0e-67

Match length 195 % identity 64

NCBI Description subtilisin-like proteinase (EC 3.4.21.-) - Arabidopsis

thaliana (fragment) >gi_757534_emb_CAA59963_ (X85974)

subtilisin-like protease [Arabidopsis thaliana]

Seq. No. 8239

Contig ID 12029 1.R1040 5'-most EST qsv701046386.hl

8240 Seq. No.

12031 1.R1040 Contig ID

5'-most EST uC-gmflminsoy075g06b1

Method BLASTX NCBI GI g2921873 BLAST score 677 5.0e-71 E value 245 Match length % identity 52

(AF047472) spleen mitotic checkpoint BUB3 [Homo sapiens] NCBI Description

>gi 2981231 gb AAC06258 (AF053304) mitotic checkpoint component Bub3 [Homo sapiens] >gi 3639060 (AF081496)

kinetochore protein BUB3 [Homo sapiens]

Seq. No. 8241

Contig ID 12038 1.R1040 5'-most EST leu701153034.h1

Method BLASTX NCBI GI g4388731 BLAST score 283 E value 2.0e-25 Match length 94

% identity 64

NCBI Description (AC006413) hypothetical protein [Arabidopsis thaliana]

Seq. No.

8242

Contig ID

12041_1.R1040

5'-most EST LIB3170-009-Q2-K1-H4

Seq. No. 8243

12049 1.R1040 Contig ID

5'-most EST LIB3040-016-Q1-E1-H2

Seq. No. 8244

12049 2.R1040 Contig ID

5'-most EST LIB3040-041-Q1-E1-G10

Seq. No. 8245

Contig ID 12063 1.R1040

5'-most EST LIB3040-041-Q1-E1-D7

Seq. No. 8246



Contig ID 12076 1.R1040 5'-most EST LIB3170-012-Q1-K1-H6

Seq. No. 8247

Contig ID 12080_1.R1040

5'-most EST LIB3051-029-Q1-K1-A1

Method BLASTX
NCBI GI g4544434
BLAST score 518
E value 2.0e-52
Match length 142
% identity 73

NCBI Description (AC006955) putative DNA-directed RNA polymerase II

[Arabidopsis thaliana]

Seq. No. 8248

Contig ID 12083_1.R1040

5'-most EST LIB3092-062-Q1-K1-C9

Method BLASTN
NCBI GI g4558521
BLAST score 44
E value 3.0e-15
Match length 152
% identity 82

NCBI Description Genomic sequence for Arabidopsis thaliana BAC T10024,

complete sequence

Seq. No. 8249

Contig ID 12088 1.R1040 5'-most EST hrw70\overline{1}059278.h1

Seq. No. 8250

Contig ID 12096 1.R1040

5'-most EST LIB3040-040-Q1-E1-H8

Seq. No. 8251

Contig ID 12106 1.R1040

5'-most EST LIB3072-056-Q1-K1-E11

Method BLASTX
NCBI GI g4335723
BLAST score 273
E value 4.0e-24
Match length 79
% identity 61

NCBI Description (AC006248) putative thioredoxin M [Arabidopsis thaliana]

Seq. No. 8252

Contig ID 12115 1.R1040

5'-most EST LIB3170-009-Q2-K1-G10

Seq. No. 8253

Contig ID 12121 1.R1040 5'-most EST seb700651777.h1

Method BLASTX
NCBI GI g4115377
BLAST score 613
E value 1.0e-63



Match length 70 % identity

(AC005967) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

12121 2.R1040 Contig ID kl1701202752.h1 5'-most EST

BLASTX Method q4115377 NCBI GI 221 BLAST score 3.0e-18 E value 55 Match length % identity

(AC005967) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

12131 1.R1040 Contig ID

uC-qmflminsoy075g05b1 5'-most EST

BLASTX Method g3004550 NCBI GI BLAST score 175 8.0e-14 E value 86 Match length % identity 51

(AC003673) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 8256

12134 1.R1040 Contig ID

LIB3040-040-Q1-E1-E5 5'-most EST

Method BLASTX q3176668 NCBI GI 809 BLAST score 2.0e-86 E value 214 Match length 86 % identity

(AC004393) Similar to ribosomal protein L17 gb X62724 from NCBI Description Hordeum vulgare. ESTs gb Z34728, gb F19974, gb_T75677 and

gb_Z33937 come from this gene. [Arabidopsis thaliana]

8257 Seq. No.

12136 1.R1040 Contig ID

LIB3170-010-Q1-J1-G9 5'-most EST

8258 Seq. No.

12141 1.R1040 Contig ID

LIB3040-040-Q1-E1-B4 5'-most EST

8259 Seq. No.

12143 1.R1040 Contig ID

LIB3049-012-Q1-E1-A7 5'-most EST

BLASTX Method q3510255 NCBI GI 251 BLAST score 4.0e-21 E value 70 Match length 73 % identity

NCBI Description (AC005310) hypothetical protein [Arabidopsis thaliana]

NCBI Description

thaliana]



```
8260
Seq. No.
                   12148 1.R1040
Contig ID
5'-most EST
                   leu701145636.hl
                   BLASTX
Method
                   q4567246
NCBI GI
                   629
BLAST score
                   2.0e-65
E value
Match length
                   180
                   70
% identity
NCBI Description (AC007070) unknown protein [Arabidopsis thaliana]
Seq. No.
                   8261
                   12149 1.R1040
Contig ID
                   zzp700833747.h1
5'-most EST
                   BLASTX
Method
                   g217909
NCBI GI
BLAST score
                   895
E value
                   1.0e-96
Match length
                   203
                   87
% identity
NCBI Description (D14044) glycolate oxidase [Cucurbita sp.]
                   8262
Seq. No.
                   12149 2.R1040
Contig ID
                   jex700905933.hl
5'-most EST
Method
                   BLASTX
NCBI GI
                   q217909
BLAST score
                   267
E value
                   1.0e-23
Match length
                   64
% identity
                   81
NCBI Description (D14044) glycolate oxidase [Cucurbita sp.]
                   8263
Seq. No.
                   12151 1.R1040
Contig ID
5'-most EST
                   LIB3040-040-Q1-E1-B1
Seq. No.
                   8264
Contiq ID
                   12165 1.R1040
5'-most EST
                   LIB3139-059-P1-N1-E9
Seq. No.
                   8265
                   12165 2.R1040
Contig ID
5'-most EST
                   LIB3106-026-Q1-K1-E7
Seq. No.
                   8266
                   12169 1.R1040
Contig ID
5'-most EST
                   LIB3040-040-Q1-E1-A9
Method
                   BLASTX
NCBI GI
                   g3242729
BLAST score
                   907
                   3.0e-98
E value
Match length
                   173
% identity
```

(AC003040) curly leaf protein (polycomb-group) [Arabidopsis



```
Seq. No.
                  8267
                  12171 1.R1040
Contig ID
5'-most EST
                  LIB3138-083-P1-N1-C10
Method
                  BLASTX
NCBI GI
                  g4455361
BLAST score
                  368
E value
                  3.0e-60
Match length
                  185
% identity
                  61
NCBI Description (AL035524) putative protein [Arabidopsis thaliana]
                  8268
Seq. No.
                  12175 1.R1040
Contig ID
5'-most EST
                  LIB3040-039-Q1-E1-F7
                  BLASTN
Method
NCBI GI
                  q166409
BLAST score
                  126
                  3.0e-64
E value
                  292
Match length
% identity
NCBI Description Alfalfa nucleic acid binding protein (alfin-1) mRNA,
                  partial cds
                  8269
Seq. No.
Contig ID
                  12181 1.R1040
5'-most EST
                  LIB3093-034-Q1-K1-D3
Method
                  BLASTX
                  q1173027
NCBI GI
BLAST score
                  442
                  1.0e-43
E value
Match length
                  120
% identity
                  73
NCBI Description 60S RIBOSOMAL PROTEIN L31 >gi_915313 (U23784) ribosomal
                  protein L31 [Nicotiana glutinosa]
Seq. No.
                  8270
                  12184 1.R1040
Contig ID
                  LIB3040-039-Q1-E1-G7
5'-most EST
Method
                  BLASTX
                  g2144183
NCBI GI
BLAST score
                  354
                  3.0e-33
E value
Match length
                  103
                  64
% identity
                  DNA-directed RNA polymerase (EC 2.7.7.6) chain III -
NCBI Description
                  Ecotype Columbia >gi 1184686 (U35049) RNA polymerase I(A)
                  and III(C) 14 kDa subunit [Arabidopsis thaliana]
```

>gi_1184688 (U35050) Arabidopsis thaliana RNA polymerase I(A) and III(C) 14 kDa subunit [Arabidopsis thaliana] >gi_3980382 (AC004561) RNA polymerase I(A) and III(C) 14 kDa subunit (AtpRAC14) [Arabidopsis thaliana]

kDa subunit (AtRPAC14) [Arabidopsis thaliana]

Seq. No. 8271

Contig ID 12192_1.R1040

5'-most EST LIB3170-054-Q1-K1-C2



Seq. No. 8272

Contig ID 12192 2.R1040

5'-most EST LIB3049-049-Q1-E1-G4

Seq. No. 8273

Contig ID 12200_1.R1040

5'-most EST LIB3050-024-Q1-K1-E9

Seq. No. 8274

Contig ID 12200 2.R1040 5'-most EST rca701000961.h1

Method BLASTX
NCBI GI g3242709
BLAST score 306
E value 1.0e-53
Match length 254
% identity 48

NCBI Description (AC003040) putative guanine nucleotide-binding protein

[Arabidopsis thaliana]

Seq. No. 8275

Contig ID 12202 1.R1040 5'-most EST leu701148856.h1

Seq. No. 8276

Contig ID 12207_1.R1040 5'-most EST kmv700741774.h1

Method BLASTX
NCBI GI g3041738
BLAST score 512
E value 6.0e-52
Match length 140
% identity 69

NCBI Description T-COMPLEX PROTEIN 1, ETA SUBUNIT (TCP-1-ETA) (CCT-ETA)

(HIV-1 NEF INTERACTING PROTEIN) >gi_2559010 (AF026292) chaperonin containing t-complex polypeptide 1, eta subunit;

CCT-eta [Homo sapiens]

Seq. No. 8277

Contig ID 12217_1.R1040 5'-most EST trc700567909.h1

Method BLASTX
NCBI GI g4539403
BLAST score 360
E value 4.0e-34
Match length 144
% identity 51

NCBI Description (AL049524) putative protein [Arabidopsis thaliana]

Seq. No. 8278

Contig ID 12222_1.R1040

5'-most EST LIB3051-027-Q1-K1-E7

Seq. No. 8279

Contig ID 12231_1.R1040 5'-most EST jex700907994.h1

Method BLASTX

E value

Match length

% identity

107

56



```
q3212871
NCBI GI
                   474
BLAST score
                   3.0e-89
E value
                  186
Match length
                   87
% identity
                   (AC004005) putative translation initiation factor
NCBI Description
                   [Arabidopsis thaliana]
                   8280
Seq. No.
                   12231 2.R1040
Contig ID
                   LIB3040-038-Q1-E1-H8
5'-most EST
                   BLASTX
Method
                   g3212871
NCBI GI
                   264
BLAST score
                   6.0e-23
E value
                   55
Match length
                   85
% identity
                   (AC004005) putative translation initiation factor
NCBI Description
                   [Arabidopsis thaliana]
                   8281
Seq. No.
                   12231 3.R1040
Contig ID
                   dpv701097109.h1
5'-most EST
                   BLASTX
Method
                   q3212871
NCBI GI
                   249
BLAST score
                   3.0e-21
E value
Match length
                   53
% identity
                   (AC004005) putative translation initiation factor
NCBI Description
                   [Arabidopsis thaliana]
                   8282
Seq. No.
                   12232 1.R1040
Contig ID
                   LIB3051-039-Q1-K1-G3
 5'-most EST
                   BLASTX
Method
                   g2832645
NCBI GI
BLAST score
                   164
                   3.0e-11
 E value
                   55
 Match length
 % identity
                   73
                   (AL021710) hypothetical protein [Arabidopsis thaliana]
 NCBI Description
                   8283
 Seq. No.
                   12232 2.R1040
 Contig ID
                   LIB3106-039-Q1-K1-G7
 5'-most EST
 Seq. No.
                   8284
                   12243 1.R1040
 Contig ID
                   LIB3170-010-Q1-K1-F12
 5'-most EST
                   BLASTX
 Method
                    g4454043
 NCBI GI
                    314
 BLAST score
                    4.0e-29
```

1507

NCBI Description (AL035394) putative receptor kinase [Arabidopsis thaliana]



```
8285
Seq. No.
                  12244 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy097a11b1
Method
                  BLASTX
NCBI GI
                  g2190419
BLAST score
                  480
E value
                  5.0e-48
Match length
                  150
% identity
                  58
NCBI Description (Y13632) dem [Lycopersicon esculentum]
                  8286
Seq. No.
                  12255 1.R1040
Contig ID
5'-most EST
                  LIB3049-010-Q1-E1-D10
Method
                  BLASTN
NCBI GI
                  g4159703
BLAST score
                  57
E value
                  7.0e-23
                  325
Match length
% identity
                  85
NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:
                  K5F14, complete sequence [Arabidopsis thaliana]
                   8287
Seq. No.
Contig ID
                  12255 2.R1040
5'-most EST
                  LIB3092-022-Q1-K1-C10
Method
                  BLASTN
NCBI GI
                  g4159703
BLAST score
                  57
                   6.0e-23
E value
Match length
                  151
                   90
% identity
NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:
                   K5F14, complete sequence [Arabidopsis thaliana]
Seq. No.
                   8288
                   12255 3.R1040
Contig ID
5'-most EST
                   zhf700965132.h1
Method
                   BLASTX
NCBI GI
                   g2244797
BLAST score
                   284
                   4.0e-25
E value
Match length
                   84
% identity
NCBI Description (Z97336) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   8289
```

Contig ID 12255_4.R1040 5'-most EST uC-gmropic006d07b1

Seq. No. 8290

Contig ID 12255 5.R1040

5'-most EST jC-gmle01810068b06a1

Seq. No. 8291

Contig ID 12255_6.R1040



```
jC-gmro02910002b01a1
5'-most EST
                  8292
Seq. No.
                  12255 8.R1040
Contig ID
5'-most EST
                  pxt700942169.hl
                  8293
Seq. No.
                  12257 1.R1040
Contig ID
                  LIB3170-009-Q2-K2-F7
5'-most EST
                  8294
Seq. No.
                  12267 1.R1040
Contig ID
                  LIB3040-038-Q1-E1-D6
5'-most EST
                  8295
Seq. No.
                  12268 1.R1040
Contig ID
5'-most EST
                  LIB3170-009-Q2-K1-F9
                  8296
Seq. No.
                  12270 1.R1040
Contig ID
5'-most EST
                  sat701012249.hl
Method
                  BLASTN
                  q14162
NCBI GI
                  41
BLAST score
                  2.0e-13
E value
Match length
                   53
% identity
NCBI Description P.hybrida mitochondria genes rps19, Yrps3 and Yrpl16
Seq. No.
                   8297
                   12271 1.R1040
Contig ID
5'-most EST
                   leu701147506.hl
                   BLASTX
Method
                   g169459
NCBI GI
                   239
BLAST score
                   4.0e-20
E value
Match length
                   93
                   49
% identity
                  (M18538) pop3 peptide [Populus balsamifera subsp.
NCBI Description
                   trichocarpa X Populus deltoides]
                   8298
Seq. No.
                   12274 1.R1040
Contig ID
                   LIB3040-038-Q1-E1-A12
5'-most EST
                   BLASTX
Method
                   g4006872
NCBI GI
                   301
BLAST score
                   1.0e-27
E value
Match length
                   65
                   85
% identity
                   (Z99707) methionyl aminopeptidase-like protein [Arabidopsis
NCBI Description
```

Seq. No.

8299

Contig ID

12277 1.R1040

thaliana]

5'-most EST

LIB3072-021-Q1-E1-F5

Method

BLASTX

g418507

NCBI GI g418507 BLAST score 342 E value 6.0e-32 Match length 156 % identity 46

NCBI Description S-ADENOSYLMETHIONINE: 2-DEMETHYLMENAQUINONE

METHYLTRANSFERASE >gi_541097_pir__S40872 hypothetical protein f161 - Escherichia coli >gi_305032 (L19201) ORF_f161 [Escherichia coli] >gi_1336002 (U56082)

S-adenosylmethionine:2-demethylmenaquinone

methyltransferase [Escherichia coli] >gi_1790364 (AE000467)

menaquinone biosynthesis, unknown [Escherichia coli]

Seq. No. 8300

Contig ID 12277_2.R1040

5'-most EST LIB31 $\frac{1}{3}$ 8-064-Q1-N1-B10

Method BLASTX
NCBI GI g418507
BLAST score 165
E value 2.0e-11
Match length 53
% identity 58

NCBI Description S-ADENOSYLMETHIONINE: 2-DEMETHYLMENAQUINONE

METHYLTRANSFERASE >gi_541097_pir__S40872 hypothetical protein f161 - Escherichia coli >gi_305032 (L19201) ORF_f161 [Escherichia coli] >gi_1336002 (U56082)

S-adenosylmethionine:2-demethylmenaquinone

methyltransferase [Escherichia coli] >gi_1790364 (AE000467)

menaquinone biosynthesis, unknown [Escherichia coli]

Seq. No. 8301

Contig ID 12279_1.R1040

5'-most EST LIB3087-012-Q1-K1-E5

Method BLASTX
NCBI GI g2160182
BLAST score 293
E value 4.0e-26
Match length 155
% identity 44

NCBI Description (AC000132) ESTs gb_ATTS1236, gb_T43334, gb_N97019, gb_AA395203

come from this gene. [Arabidopsis thaliana]

Seq. No. 8302

Contig ID 12279 2.R1040

5'-most EST LIB3107-057-Q1-K1-F1

Seq. No. 8303

Contig ID 12282_1.R1040

5'-most EST LIB3074-011-Q1-E1-H7

Method BLASTX
NCBI GI g466160
BLAST score 372
E value 2.0e-35
Match length 83
% identity 86

NCBI Description HYPOTHETICAL 9.8 KD PROTEIN ZK652.3 IN CHROMOSOME III

>gi 630771 pir S44903 ZK652.3 protein - Caenorhabditis



elegans >g 289769 (L14429) putative [Caenorhabditis elegans]

Seq. No. 8304

Contig ID 12282_2.R1040

5'-most EST LIB3049-025-Q1-E1-C1

Method BLASTX
NCBI GI g466160
BLAST score 372
E value 1.0e-35
Match length 83
% identity 86

NCBI Description HYPOTHETICAL 9.8 KD PROTEIN ZK652.3 IN CHROMOSOME III

>gi_630771_pir__S44903 ZK652.3 protein - Caenorhabditis
elegans >gi 289769 (L14429) putative [Caenorhabditis

elegans]

Seq. No. 8305

Contig ID 12288 1.R1040 5'-most EST sat701013128.h1

Method BLASTX
NCBI GI g1709798
BLAST score 1888
E value 0.0e+00
Match length 388
% identity 96

NCBI Description 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG >gi 1155334

(U43398) POTATP1 [Solanum tuberosum]

Seq. No. 8306

Contig ID 12288_2.R1040 5'-most EST fua701042085.h1

Method BLASTX
NCBI GI g1709798
BLAST score 453
E value 3.0e-45
Match length 102
% identity 90

NCBI Description 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG >gi_1155334

(U43398) POTATP1 [Solanum tuberosum]

Seq. No. 8307

Contig ID 12291 1.R1040

5'-most EST LIB3040-037-Q1-E1-F8

Seq. No. 8308

Contig ID 12302 1.R1040 5'-most EST leu701151648.h1

Method BLASTX
NCBI GI g1350548
BLAST score 350
E value 1.0e-32
Match length 115
% identity 57

NCBI Description (L47609) heat shock-like protein [Picea glauca]



```
12302 2.R1040
Contig ID
                   k11701204314.h2
5'-most EST
                   BLASTX
Method
                   q1350548
NCBI GI
BLAST score
                   334
E value
                   8.0e-31
                   108
Match length
                   57
% identity
NCBI Description (L47609) heat shock-like protein [Picea glauca]
Seq. No.
                   8310
Contig ID
                   12302 3.R1040
5'-most EST
                   LIB3170-029-Q1-K1-G12
                   BLASTX
Method
                   q1350548
NCBI GI
BLAST score
                   281
E value
                   4.0e-25
Match length
                   85
% identity
                   (L47609) heat shock-like protein [Picea glauca]
NCBI Description
                   8311
Seq. No.
                   12302 4.R1040
Contig ID
5'-most EST
                   LIB3138-011-Q1-N2-B5
Seq. No.
                   8312
Contig ID
                   12302 5.R1040
5'-most EST
                   LIB3093-001-Q1-K1-F10
                   BLASTX
Method
                   g1350548
NCBI GI
BLAST score
                   277
E value
                   4.0e-25
                   93
Match length
                   59
% identity
                   (L47609) heat shock-like protein [Picea glauca]
NCBI Description
                   8313
Seq. No.
                   12305 1.R1040
Contig ID
                   LIB3040-037-Q1-E1-H5
5'-most EST
                   8314
Seq. No.
                   12307 1.R1040
Contig ID
                   epx70\overline{1}110127.h1
5'-most EST
                   BLASTX
Method
                   g3063450
NCBI GI
                   204
BLAST score
                   2.0e-23
E value
                   108
Match length
                   62
% identity
                   (AC003981) F22013.12 [Arabidopsis thaliana]
NCBI Description
                   8315
Seq. No.
                   12310 1.R1040
Contig ID
                   qsv701045922.h1
5'-most EST
```

1512

BLASTN

g758643 291

Method NCBI GI

BLAST score



E value 1.0e-162 Match length 487 % identity 90

NCBI Description P.sativum mRNA for nucleoside diphosphate kinase II

Seq. No. 8316

Contig ID 12320_1.R1040 5'-most EST txt700733277.h1

Method BLASTN
NCBI GI g2921322
BLAST score 191
E value 1.0e-103
Match length 199
% identity 99

NCBI Description Glycine max beta-1,3-glucanase 7 (SGlu7) gene, partial cds

Seq. No. 8317

Contig ID 12335 1.R1040

5'-most EST LIB3040-037-Q1-E1-C4

Seq. No. 8318

Contig ID 12336 1.R1040

5'-most EST LIB3040-037-Q1-E1-A4

Seq. No. 8319

Contig ID 12337_1.R1040 5'-most EST epx701107811.h1

Method BLASTX
NCBI GI g3810835
BLAST score 219
E value 2.0e-17
Match length 132
% identity 40

NCBI Description (AL032684) putative RNA-binding protein

[Schizosaccharomyces pombe]

Seq. No. 8320

Contig ID 12337_2.R1040 5'-most EST sat701011382.h1

Seq. No. 8321

Contig ID 12341_1.R1040

5'-most EST LIB3040-036-Q1-E1-G8

Seq. No. 8322

Contig ID 12343_1.R1040

5'-most EST LIB3040-017-Q1-E1-A5

Seq. No. 8323

Contig ID 12351 1.R1040 5'-most EST uxk700670327.h1

Method BLASTN
NCBI GI g2924257
BLAST score 368
E value 0.0e+00
Match length 866
% identity 88



NCBI Description Tobacco chloroplast genome DNA

Seq. No. 8324

Contig ID 12351 2.R1040

5'-most EST LIB3107-017-Q1-K1-B8

Method BLASTN
NCBI GI g12283
BLAST score 402
E value 0.0e+00
Match length 900
% identity 90

NCBI Description Spinach plastid psbB operon with genes for 10 kD

phosphoprotein associated with photosystem II (psbH), apocytochrome b6 and subunit 4 (petD) of cytochrome b6f

complex

Seq. No. 8325

Contig ID 12351_3.R1040 5'-most EST rca700998107.h1

Method BLASTN
NCBI GI g2924257
BLAST score 241
E value 1.0e-132
Match length 964
% identity 89

NCBI Description Tobacco chloroplast genome DNA

Seq. No. 8326

Contig ID 12351 4.R1040

5'-most EST LIB3170-040-Q1-J1-D4

Method BLASTN
NCBI GI g2143322
BLAST score 157
E value 1.0e-82
Match length 305
% identity 91

NCBI Description P.deltoides chloroplast DNA for psbB operon

Seq. No. 8327

Contig ID 12351_6.R1040 5'-most EST rlr700898879.h1

Method BLASTN
NCBI GI g12150
BLAST score 163
E value 2.0e-86
Match length 223
% identity 93

NCBI Description Pea chloroplast gene for 15.2 kDA polypeptide from

cytochrome b-f complex

Seq. No. 8328

Contig ID 12352 1.R1040

5'-most EST LIB3040-036-Q1-E1-H9

Seq. No. 8329

Contig ID 12353_1.R1040 5'-most EST ncj700982094.h1



```
8330
Seq. No.
                   12353 2.R1040
Contig ID
                  hrw701061763.hl
5'-most EST
Seq. No.
                   8331
                   12354 1.R1040
Contig ID
                   uC-gm\overline{f}lminsoy042d02b1
5'-most EST
                   BLASTX
Method
                   g4539352
NCBI GI
BLAST score
                   200
                   2.0e-16
E value
                   156
Match length
                   37
% identity
                   (AL035539) putative protein [Arabidopsis thaliana]
NCBI Description
                   8332
Seq. No.
                   12354 2.R1040
Contig ID
                   LIB3051-038-Q1-K1-C10
5'-most EST
                   BLASTX
Method
                   q4539352
NCBI GI
BLAST score
                   626
                   4.0e-65
E value
                   240
Match length
% identity
                   (AL035539) putative protein [Arabidopsis thaliana]
NCBI Description
                   8333
Seq. No.
                   12356 1.R1040
Contig ID
                   LIB3170-009-Q2-K1-E12
5'-most EST
                   BLASTX
Method
                   q4220462
NCBI GI
                   601
BLAST score
                   1.0e-62
E value
                   127
Match length
                   88
% identity
                   (AC006216) Strong similarity to gb Z50851 HD-zip (athb-8)
NCBI Description
                   gene from Arabidopsis thaliana containing Homeobox PF_00046
                   and bZIP PF_00170 domains. [Arabidopsis thaliana]
                   8334
Seq. No.
                   12357 1.R1040
Contig ID
                   LIB3107-076-Q1-K1-D4
 5'-most EST
                   BLASTX
Method
                   g4093155
NCBI GI
                   575
BLAST score
                   5.0e-59
E value
                   196
Match length
 % identity
                   (AF088281) phytochrome-associated protein 1 [Arabidopsis
NCBI Description
                   thaliana]
```

Contig ID 5'-most EST 12357 3.R1040 zhf700953009.h1

Method BLASTX q4093155 NCBI GI

```
BLAST score
                  1.0e-11
E value
Match length
                  48
% identity
                  (AF088281) phytochrome-associated protein 1 [Arabidopsis
NCBI Description
                  8336
Seq. No.
                  12360 1.R1040
Contig ID
                  LIB3040-036-Q1-E1-E7
5'-most EST
Method
                  BLASTX
NCBI GI
                  q3281868
BLAST score
                  532
E value
                   3.0e-54
Match length
                  133
% identity
                   74
                  (AL031004) putative protein [Arabidopsis thaliana]
NCBI Description
                   8337
Seq. No.
                   12365 1.R1040
Contig ID
5'-most EST
                   zzp700835138.h1
Method
                   BLASTX
NCBI GI
                   g2582971
BLAST score
                   218
                   4.0e-18
E value
Match length
                   76
% identity
                   59
                  (D83711) TKRP125 [Nicotiana tabacum]
NCBI Description
                   8338
Seq. No.
                   12375 1.R1040
Contig ID
5'-most EST
                   jC-qmf102220106e09a1
                   8339
Seq. No.
                   12380 1.R1040
Contig ID
                   taw700660389.hl
5'-most EST
                   8340
Seq. No.
                   12383 1.R1040
Contig ID
                   LIB3107-016-Q1-K1-A2
5'-most EST
                   BLASTX
Method
                   g2983600
NCBI GI
BLAST score
                   1036
                   1.0e-113
E value
                   399
Match length
                   52
% identity
                  (AE000725) argininosuccinate synthase [Aquifex aeolicus]
NCBI Description
```

8341 Seq. No. Contig ID

12383 3.R1040 5'-most EST pmv700894244.h1 BLASTX Method

NCBI GI q2983600 223 BLAST score E value 5.0e-18 Match length 93 % identity 52



```
(AE000725) argininosuccinate synthase [Aquifex aeolicus]
NCBI Description
                  8342
Seq. No.
Contig ID
                  12383 4.R1040
5'-most EST
                  gsv701044237.hl
Seq. No.
                  8343
                  12387 1.R1040
Contig ID
                  LIB3139-048-P1-N1-A2
5'-most EST
Method
                  BLASTN
                  g1694899
NCBI GI
BLAST score
                  391
                  0.0e + 00
E value
Match length
                  583
                   92
% identity
NCBI Description P.sativum mRNA for Cop1 protein
Seq. No.
                   8344
                   12388 1.R1040
Contig ID
5'-most EST
                   g5057689
Method
                   BLASTX
                   q3193320
NCBI GI
BLAST score
                   154
                   7.0e-10
E value
Match length
                   78
% identity
                   38
                   (AF069299) contains similarity to the subtilase family of
NCBI Description
                   serine proteases (Pfam: subtilase.hmm, score: 47.57);
                   strong similarity to Cucumis melo (muskmelon) cucumisin
                   (GB:D32206) [Arabidopsis thaliana]
                   8345
Seq. No.
                   12391 1.R1040
Contig ID
                   LIB3170-010-Q1-J1-E7
5'-most EST
Seq. No.
                   8346
                   12393 1.R1040
Contig ID
                   zhf700963357.h1
5'-most EST
                   8347
Seq. No.
                   12395 1.R1040
Contig ID
                   LIB3040-036-Q1-E1-E11
5'-most EST
                   8348
Seq. No.
                   12406 1.R1040
Contig ID
5'-most EST
                   LIB3040-036-Q1-E1-B10
Seq. No.
                   8349
                   12407 1.R1040
Contig ID
                   LIB3040-036-Q1-E1-B11
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2245118
BLAST score
                   695
                   5.0e-73
E value
Match length
                   153
```

NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]

79

% identity



Contig ID 12407_2.R1040 5'-most EST taw700657218.h1

Method BLASTX
NCBI GI g2245118
BLAST score 168
E value 1.0e-11
Match length 66
% identity 55

NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]

Seq. No.

8351

Contig ID 12410 1.R1040

5'-most EST LIB3170-010-Q1-K1-E6

Method BLASTX
NCBI GI g3510256
BLAST score 467
E value 3.0e-46
Match length 151
% identity 60

NCBI Description (AC005310) unknown protein [Arabidopsis thaliana]

Seq. No.

8352

Contig ID

12417 1.R1040

5'-most EST

LIB3040-008-Q1-E1-E11

Seq. No. 8353

Contig ID

12430 1.R1040

5'-most EST

LIB3170-009-Q2-K2-E3

Seq. No.

8354

Contig ID 5'-most EST

12438_1.R1040 pxt700946066.h1

Method BLASTX
NCBI GI g3219858
BLAST score 528
E value 1.0e-53
Match length 116
% identity 84

NCBI Description DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEPTIDE

>gi_2129724_pir__S71204 RNA polymerase II 13.6 kDa chain - Arabidopsis thaliana >gi_881501 (U28048) RNA polymerase II

13.6 kDa subunit [Arabidopsis thaliana]

Seq. No. 8355

Contig ID 12438 2.R1040

5'-most EST LIB3040-035-Q1-E1-D4

Method BLASTX
NCBI GI g3219858
BLAST score 331
E value 7.0e-31
Match length 79
% identity 80

NCBI Description DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEPTIDE

>gi_2129724_pir__S71204 RNA polymerase II 13.6 kDa chain Arabidopsis thaliana >gi_881501 (U28048) RNA polymerase II



13.6 kDa subunit [Arabidopsis thaliana]

 Seq. No.
 8356

 Contig ID
 12438_3.R1040

 5'-most EST
 LIB3050-013-Q1-E1-H7

 Method
 BLASTX

 NCBI GI
 g3219858

 BLAST score
 403

 E value
 2.0e-39

91

% identity 84 NCBI Description DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEPTIDE

>gi_2129724_pir__S71204 RNA polymerase II 13.6 kDa chain Arabidopsis thaliana >gi_881501 (U28048) RNA polymerase II

13.6 kDa subunit [Arabidopsis thaliana]

Seq. No. 8357

Match length

Contig ID 12438 4.R1040 5'-most EST bth700844772.h1

Method BLASTX
NCBI GI g3219858
BLAST score 169
E value 3.0e-12
Match length 44
% identity 73

NCBI Description DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEPTIDE

>gi_2129724_pir__S71204 RNA polymerase II 13.6 kDa chain Arabidopsis thaliana >gi_881501 (U28048) RNA polymerase II

13.6 kDa subunit [Arabidopsis thaliana]

Seq. No. 8358

Contig ID 12439 1.R1040

5'-most EST LIB3170-012-Q1-K1-E6

Seq. No. 8359

Contig ID 12443 1.R1040

5'-most EST LIB3170-011-Q1-K1-E7

Seq. No. 8360

Contig ID 12446_1.R1040 5'-most EST fua701040824.h1

Method BLASTX
NCBI GI g1834379
BLAST score 183
E value 3.0e-13
Match length 98
% identity 43

NCBI Description (Y10551) ComAB [Bacillus licheniformis]

Seq. No. 8361

Contig ID 12452_1.R1040 5'-most EST zpv700763308.h1

Method BLASTN
NCBI GI g4103986
BLAST score 109
E value 6.0e-54
Match length 282



% identity 89

NCBI Description Pisum sativum 5,10-methylenetetrahydrofolate

dehydrogenase-5,10-methenyltetrahydrofolate cyclohydrolase

mRNA, complete cds

Seq. No. 8362

Contig ID 12452_2.R1040 5'-most EST bth700847483.h1

Method BLASTN
NCBI GI g4103986
BLAST score 495
E value 0.0e+00
Match length 859
% identity 89

NCBI Description Pisum sativum 5,10-methylenetetrahydrofolate

dehydrogenase-5,10-methenyltetrahydrofolate cyclohydrolase

mRNA, complete cds

Seq. No. 8363

Contig ID 12468_1.R1040 5'-most EST uC-gmropic057h01b1

Method BLASTN
NCBI GI g2062691
BLAST score 33
E value 6.0e-09
Match length 33

% identity 61

NCBI Description Human sodium phosphate transporter (NPT4) mRNA, complete

cds

Seq. No. 8364

Contig ID 12470_1.R1040

5'-most EST LIB3170-011-Q1-K1-E4

Seq. No. 8365

Contig ID 12473_1.R1040 5'-most EST pmv700892186.h1

Seq. No. 8366

Contig ID 12477 1.R1040

5'-most EST LIB3040-034-Q1-E1-H2

Method BLASTX
NCBI GI g2204234
BLAST score 522
E value 6.0e-53
Match length 183
% identity 56

NCBI Description (Y13862) enoyl-ACP reductase [Nicotiana tabacum]

Seq. No. 8367

Contig ID 12477_2.R1040 5'-most EST ssr700560123.h1

Seq. No. 8368

Contig ID 12480 1.R1040

5'-most EST LIB3050-005-Q1-K1-D7



```
Seq. No. 8369
```

Contig ID 12483 1.R1040

5'-most EST LIB3040-035-Q1-E1-A1

Seq. No. 8370

Contig ID 12484 1.R1040

5'-most EST LIB3040-034-Q1-E1-G2

Seq. No. 8371

Contig ID 12490 1.R1040

5'-most EST LIB3170-010-Q1-K1-D2

Seq. No. 8372

Contig ID 12493 1.R1040

5'-most EST jC-gmle01810063h01a1

Method BLASTX
NCBI GI g2894599
BLAST score 258
E value 4.0e-22
Match length 71
% identity 53

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8373

Contig ID 12493 2.R1040

5'-most EST jC-gmle01810010e01a1

Method BLASTX
NCBI GI g2894599
BLAST score 296
E value 1.0e-26
Match length 100
% identity 56

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8374

Contig ID 12493_3.R1040 5'-most EST bth700848738.h1

Method BLASTX
NCBI GI g2894599
BLAST score 531
E value 6.0e-57
Match length 195
% identity 64

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Seq. No. 8375

Contig ID 12493_5.R1040

5'-most EST LIB3138-033-Q1-N1-A4

Method BLASTX
NCBI GI g2894599
BLAST score 252
E value 1.0e-21
Match length 100
% identity 53

NCBI Description (AL021889) putative protein [Arabidopsis thaliana]

Contig ID

5'-most EST



```
12500 1.R1040

    ○ Contig ID

                    LIB3094-091-Q1-K1-E7
  5'-most EST
                    BLASTX
  Method
                     g3759184
  NCBI GI
  BLAST score
                     1082
                     1.0e-118
  E value
  Match length
                     290
                     70
  % identity
  NCBI Description (AB018441) phi-1 [Nicotiana tabacum]
                     8377
  Seq. No.
                     12500 2.R1040
  Contig ID
                     zsg701122572.hl
  5'-most EST
                     BLASTX
  Method
                     g3759184
  NCBI GI
                     336
  BLAST score
                     2.0e-31
  E value
                     109
  Match length
                     61
  % identity
                     (AB018441) phi-1 [Nicotiana tabacum]
  NCBI Description
                     8378
  Seq. No.
                     12500 3.R1040
  Contig ID
                     LIB3139-084-P1-N1-G3
  5'-most EST
                     BLASTX
  Method
                     g3759184
  NCBI GI
                     504
  BLAST score
                     6.0e-51
  E value
                     163
  Match length
                     60
   % identity
                     (AB018441) phi-1 [Nicotiana tabacum]
  NCBI Description
                     8379
   Seq. No.
                     12500 6.R1040
   Contig ID
                     rlr700898648.hl
   5'-most EST
                     BLASTX
  Method
                      q3759184
   NCBI GI
   BLAST score
                      164
                      3.0e-11
   E value
                      40
   Match length
                      78
   % identity
                     (AB018441) phi-1 [Nicotiana tabacum]
   NCBI Description
                      8380
   Seq. No.
                      12500 7.R1040
   Contig ID
                      LIB3139-056-P1-N1-G1
   5'-most EST
                      BLASTX
   Method
                      g4325369
   NCBI GI
                      248
   BLAST score
                      7.0e-31
   E value
                      107
   Match length
                      63
   % identity
                     (AF128396) T3H13.3 gene product [Arabidopsis thaliana]
   NCBI Description
                      8381
   Seq. No.
                      12508 1.R1040
```

1522

LIB3040-034-Q1-E1-G11



```
Method
                  q1703091
NCBI GI
                  413
BLAST score
                  3.0e-40
E value
Match length
                  115
                  75
% identity
                  ACYL CARRIER PROTEIN, MITOCHONDRIAL PRECURSOR (ACP)
NCBI Description
                   (NADH-UBIQUINONE OXIDOREDUCTASE 9.6 KD SUBUNIT) (MTACP-1)
                  >gi_903689 (L23574) acyl carrier protein precursor
                   [Arabidopsis thaliana] >gi_3341682 (AC003672) acyl carrier
                  protein [Arabidopsis thaliana]
                   8382
Seq. No.
                   12511 1.R1040
Contig ID
                   jC-gmle01810059b09a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4262154
BLAST score
                   495
                   1.0e-49
E value
                   154
Match length
                   67
% identity
                   (AC005275) putative protein phosphatase regulatory subunit
NCBI Description
                   [Arabidopsis thaliana]
                   8383
Seq. No.
                   12518 1.R1040
Contig ID
                   LIB3049-002-Q1-E1-C1
5'-most EST
                   BLASTX
Method
                   q4335763
NCBI GI
BLAST score
                   661
                   9.0e-73
E value
                   177
Match length
% identity
                   (AC006284) unknown protein [Arabidopsis thaliana]
NCBI Description
                   8384
Seq. No.
                   12518 2.R1040
Contig ID
                   gsv701049114.h1
5'-most EST
                   BLASTX
Method
                   q4335763
NCBI GI
                   151
BLAST score
                   6.0e-10
E value
Match length
                   44
                   73
% identity
                   (AC006284) unknown protein [Arabidopsis thaliana]
NCBI Description
                   8385
Seq. No.
                   12520 1.R1040
Contig ID
                   LIB3167-010-P1-K1-D7
 5'-most EST
```

Method BLASTX
NCBI GI g1402891
BLAST score 184
E value 3.0e-13
Match length 90
% identity 49

NCBI Description (X98130) unknown [Arabidopsis thaliana]



```
Seq. No.
                  12520 2.R1040
Contig ID
5'-most EST
                  smc700750345.hl
Seq. No.
                  8387
Contig ID
                  12520 4.R1040
                  LIB3138-034-Q1-N1-B6
5'-most EST
                  BLASTX
Method
NCBI GI
                  g1402890
                   655
BLAST score
                   3.0e-68
E value
```

% identity NCBI Description (X98130) unknown [Arabidopsis thaliana]

Seq. No. 12521 1.R1040 Contig ID 5'-most EST kl1701214785.h1

8389 Seq. No.

Match length

12521 2.R1040 Contig ID LIB3049-039-Q1-E1-B4

5'-most EST

287

8390 Seq. No.

12524 1.R1040 Contig ID

LIB3074-025-Q1-E1-H7 5'-most EST BLASTX

Method g2245131 NCBI GI 902 BLAST score E value 2.0e-97 Match length 262 73 % identity

(Z97344) hypothetical protein [Arabidopsis thaliana] NCBI Description

8391 Seq. No.

12524 2.R1040 Contig ID

5'-most EST jC-gmro02910004h11a1

Method BLASTX g2245131 NCBI GI BLAST score 293 E value 4.0e-37 120 Match length % identity 73

(Z97344) hypothetical protein [Arabidopsis thaliana] NCBI Description

8392 Seq. No.

12527 1.R1040 Contig ID

LIB3040-014-Q1-E1-A7 5'-most EST

BLASTN Method g914860 NCBI GI 116 BLAST score 3.0e-58 E value 369 Match length 83 % identity

NCBI Description M.varia mRNA for mitotic cyclin



Contig ID 12532_1.R1040

5'-most EST LIB3040-029-Q1-E1-F9

Method BLASTN
NCBI GI g4206101
BLAST score 152
E value 1.0e-79
Match length 176
% identity 97

NCBI Description Glycine max retroelement diaspora gag-pol polyprotein

(gag-pol) pseudogene, partial sequence

Seq. No. 8394

Contig ID 12541_1.R1040 5'-most EST uC-gmropic026d06b1

Method BLASTX
NCBI GI g464444
BLAST score 1087
E value 1.0e-119
Match length 238
% identity 87

NCBI Description PROTEASOME, 30 KD SUBUNIT (MULTICATALYTIC ENDOPEPTIDASE

COMPLEX 30 KD SUBUNIT) >gi_541889_pir_S39900 proteasome -

Arabidopsis thaliana $>gi_1\overline{16}6830 \text{ (M}984\overline{95})$ proteasome

[Arabidopsis thaliana]

Seq. No. 8395

Contig ID 12541_2.R1040

5'-most EST uC-gmflminsoy093c02b1

Method BLASTX
NCBI GI g4115377
BLAST score 1977
E value 0.0e+00
Match length 489
% identity 76

NCBI Description (AC005967) unknown protein [Arabidopsis thaliana]

Seq. No. 8396

Contig ID 12541 3.R1040

5'-most EST LIB3170-058-Q1-K1-E5

Method BLASTX
NCBI GI g464444
BLAST score 407
E value 2.0e-39
Match length 100
% identity 76

NCBI Description PROTEASOME, 30 KD SUBUNIT (MULTICATALYTIC ENDOPEPTIDASE

COMPLEX 30 KD SUBUNIT) >gi_541889_pir__S39900 proteasome -

Arabidopsis thaliana $>gi_1\overline{6}6830 \text{ (M98495)}$ proteasome

[Arabidopsis thaliana]

Seq. No. 8397

Contig ID 12541_4.R1040

5'-most EST jC-gmf102220054e09a1

Method BLASTN
NCBI GI g3421091
BLAST score 58
E value 8.0e-24



146 Match length 85 % identity

Arabidopsis thaliana 20S proteasome subunit PAF1 (PAF1) NCBI Description

mRNA, complete cds

Seq. No.

8398

Contig ID

12553 1.R1040

5'-most EST

LIB3073-023-Q1-K1-A7

Method NCBI GI BLASTX q3914555

BLAST score E value

337 1.0e-31

Match length % identity

86 85

NCBI Description PUTATIVE RIBOSOME-BINDING FACTOR A, CHLOROPLAST PRECURSOR >gi_3096942_emb_CAA18852.1_ (AL023094) putative protein

[Arabidopsis thaliana]

Seq. No.

8399

Contig ID

12555 1.R1040

5'-most EST

LIB3040-033-Q1-E1-F3

Method NCBI GI BLASTX g4262232

BLAST score

532

E value

5.0e-54 212

Match length % identity

48

NCBI Description

(AC006200) putative ribosomal protein L7 [Arabidopsis

thaliana]

Seq. No.

8400

Contig ID

12555 3.R1040

5'-most EST

LIB3049-026-Q1-E1-G7

Seq. No.

8401

Contig ID 5'-most EST 12555 4.R1040 fua701042636.hl

Seq. No.

8402

Contig ID

12561 1.R1040

5'-most EST

LIB3040-033-Q1-E1-G12

Seq. No.

8403

Contig ID

12565 1.R1040

5'-most EST

LIB3139-045-P1-N1-A6

Seq. No.

8404

Contig ID 5'-most EST 12576 1.R1040 sat701014478.h1

Method

BLASTN

NCBI GI BLAST score g2502084

E value

219

Match length

1.0e-119 431

% identity

88

NCBI Description Vigna radiata adenosine triphosphatase mRNA, partial cds



Contig ID 12576_2.R1040

5'-most EST LIB3040-033-Q1-E1-D2

Method BLASTN
NCBI GI g2502084
BLAST score 223
E value 1.0e-122
Match length 524
% identity 87

NCBI Description Vigna radiata adenosine triphosphatase mRNA, partial cds

Seq. No. 8406

Contig ID 12578_1.R1040

5'-most EST LIB3040-033-Q1-E1-D8

Method BLASTX
NCBI GI g3252868
BLAST score 978
E value 1.0e-106
Match length 332
% identity 57

NCBI Description (AF033536) putative zinc transporter [Arabidopsis thaliana]

Seq. No. 8407

Contig ID 12586_1.R1040 5'-most EST epx701109309.h1

Method BLASTX
NCBI GI g3335335
BLAST score 414
E value 2.0e-40
Match length 139
% identity 60

NCBI Description (AC004512) ESTs gb_F14113 and gb_T42122 come from this

region. [Arabidopsis thaliana]

Seq. No. 8408

Contig ID 12586_2.R1040

5'-most EST LIB3040-032-Q1-E1-H6

Method BLASTX
NCBI GI g3335335
BLAST score 270
E value 1.0e-23
Match length 107
% identity 54

NCBI Description (AC004512) ESTs gb_F14113 and gb_T42122 come from this

region. [Arabidopsis thaliana]

Seq. No. 8409

Contig ID 12590_1.R1040

5'-most EST LIB3139-028-P1-N1-E11

Seq. No. 8410

Contig ID 12590_2.R1040

5'-most EST LIB3055-010-Q1-N1-B6

Seq. No. 8411

Contig ID 12590_3.R1040 5'-most EST wrg700789462.h2



Contig ID 12590_5.R1040 5'-most EST zpv700758707.h1

Seq. No. 8413

Contig ID 12604_1.R1040

5'-most EST LIB3170-009-Q2-J1-D6

Seq. No. 8414

Contig ID 12608_1.R1040

5'-most EST LIB3170-011-Q1-J1-C9

Seq. No. 8415

Contig ID 12609_1.R1040 5'-most EST bth700849741.h1

Seq. No. 8416

Contig ID 12613_1.R1040 5'-most EST vwf700674168.h1

Seq. No. 8417

Contig ID 12613 2.R1040 5'-most EST vwf700673719.h1

Seq. No. 8418

Contig ID 12618_1.R1040

5'-most EST LIB3040-032-Q1-E1-F6

Seq. No. 8419

Contig ID 12621_1.R1040

5'-most EST LIB3040-032-Q1-E1-D8

Seq. No. 8420

Contig ID 12633 1.R1040

5'-most EST LIB3040-032-Q1-E1-F4

Seq. No. 8421

Contig ID 12648_1.R1040

5'-most EST LIB3040-020-Q1-E1-B5

Method BLASTX
NCBI GI g2914702
BLAST score 322
E value 2.0e-29
Match length 222
% identity 50

NCBI Description (AC003974) unknown protein [Arabidopsis thaliana]

Seq. No. 8422

Contig ID 12652_1.R1040

5'-most EST LIB3051-048-Q1-K1-D8

Seq. No. 8423

Contig ID 12671_1.R1040

5'-most EST LIB3170-012-Q1-K1-C1



12679 1.R1040 Contig ID leu701155621.h1 5'-most EST Method BLASTX NCBI GI g4417304 BLAST score 534 E value 3.0e-54 Match length 183 % identity 54

(AC006446) putative beta-1,4-mannosyl-glycoprotein NCBI Description beta-1,4-N-acetylglucosaminyltransferase [Arabidopsis

thaliana]

Seq. No.

8425 12679 2.R1040 Contig ID

5'-most EST

LIB3056-013-Q1-N1-D7

BLASTX Method NCBI GI q4417304 BLAST score 256 9.0e-22 E value 87 Match length

% identity

NCBI Description (AC006446) putative beta-1,4-mannosyl-glycoprotein

beta-1,4-N-acetylglucosaminyltransferase [Arabidopsis

thaliana]

Seq. No. 8426

Contig ID 12680 1.R1040 k11701210069.h1 5'-most EST

Seq. No. 8427

12680 2.R1040 Contig ID 5'-most EST sat701005920.h1

8428 Seq. No.

12680 3.R1040 Contig ID 5'-most EST trc700566611.h1

Seq. No. 8429

12690 1.R1040 Contig ID

5'-most EST LIB3170-009-Q2-K1-B11

8430 Seq. No.

12697 1.R1040 Contig ID

5'-most EST LIB3040-031-Q1-E2-E6

8431 Seq. No.

12698 1.R1040 Contig ID

5'-most EST LIB3040-031-Q1-E2-E7

Method BLASTX g2656003 NCBI GI BLAST score 333 E value 3.0e-31 Match length 91 % identity 73

(Z98980) hypothetical protein [Schizosaccharomyces pombe] NCBI Description



```
12708 1.R1040
Contig ID
                  LIB3094-025-Q1-K1-E12
5'-most EST
                  8433
Seq. No.
                  12708 2.R1040
Contig ID
                  LIB3040-031-Q1-E2-C1
5'-most EST
                  BLASTX
Method
                  q2865623
NCBI GI
                  668
BLAST score
E value
                  5.0e-70
                  169
Match length
                   76
% identity
                   (AF045286)
NCBI Description
                   [Arabidopsis thaliana]
```

8434

12708 3.R1040

GDP-4-keto-6-deoxy-D-mannose-3,5-epimerase-4-reductase

Contig ID LIB3087-006-Q1-K1-F4 5'-most EST BLASTX Method NCBI GI g2865623 BLAST score 188 4.0e-14 E value Match length 46 % identity 83

Seq. No.

NCBI Description (AF045286)

GDP-4-keto-6-deoxy-D-mannose-3,5-epimerase-4-reductase

[Arabidopsis thaliana]

8435 Seq. No.

12713 1.R1040 Contig ID $6HA - 0\overline{2} - Q1 - B1 - E5$ 5'-most EST

BLASTX Method g134794 NCBI GI 513 BLAST score 9.0e-52 E value 173 Match length 57 % identity

MICROSOMAL SIGNAL PEPTIDASE 18 KD SUBUNIT (SPC18) NCBI Description

>gi_108137_pir__A35309 signal peptidase (EC 3.4.99.-) 18K
chain - dog >gi_164082 (J05466) microsomal signal peptidase

complex [Canis familiaris] >gi 4335939_gb_AAD17526_ (AF061737) microsomal signal peptidase [Homo sapiens]

8436 Seq. No.

12715 1.R1040 Contig ID ncj700987568.h1 5'-most EST

BLASTN Method NCBI GI q2351064 47 BLAST score E value 3.0e-17 187 Match length

% identity Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description

MDJ22, complete sequence [Arabidopsis thaliana]

E value

Match length

1.0e-90 410

12716 1.R1040 Contig ID uaw700664678.hl 5'-most EST Method BLASTX NCBI GI q4468991 BLAST score 183 2.0e-13 E value 104 Match length 39 % identity NCBI Description (AL035605) ribosomal-like protein [Arabidopsis thaliana] Seq. No. 8438 12723 1.R1040 Contig ID LIB3040-031-Q1-E1-H4 5'-most EST BLASTN Method q971167 NCBI GI BLAST score 114 3.0e-57 E value 254 Match length 86 % identity Pisum sativum Wando ornithine carbamoyltransferase mRNA, NCBI Description complete cds 8439 Seq. No. 12728 1.R1040 Contig ID ssr700556271.hl 5'-most EST BLASTX Method g4455213 NCBI GI BLAST score 1168 1.0e-128 E value 258 Match length % identity (AL035440) glutamine amidotransferase/cyclase [Arabidopsis NCBI Description thaliana] 8440 Seq. No. 12730_1.R1040 Contig ID LIB3040-031-Q1-E2-A7 5'-most EST BLASTX Method g4455293 NCBI GI 267 BLAST score 3.0e-23E value 61 Match length 87 % identity (AL035528) putative protein [Arabidopsis thaliana] NCBI Description Seq. No. 8441 12730 2.R1040 Contig ID $k1170\overline{1}213870.h1$ 5'-most EST 8442 Seq. No. 12750 1.R1040 Contig ID LIB3049-055-Q1-E1-F3 5'-most EST BLASTN Method q488564 NCBI GI 170 BLAST score



% identity NCBI Description

Medicago sativa cultivar Chief histone H3.1 gene, 5' region and complete cds

Seq. No.

8443

Contig ID 5'-most EST 12752 1.R1040 epx701104867.h1

Seq. No.

8444

Contig ID

12752 2.R1040

5'-most EST

LIB3050-025-Q1-K1-D2

Seq. No.

8445

Contig ID

12752 3.R1040

5'-most EST

LIB3093-037-Q1-K1-B6

Seq. No.

8446

Contig ID

12759 1.R1040

5'-most EST

LIB3040-031-Q1-E1-C9

Seq. No.

8447

Contig ID

12769 1.R1040

5'-most EST

LIB3040-030-Q1-E1-G5

Method NCBI GI BLASTX g2829751

BLAST score

206

E value

5.0e-16

Match length % identity

114

NCBI Description

MACROPHAGE MIGRATION INHIBITORY FACTOR HOMOLOG (BMMIF)

>gi_1850559 (U88035) macrophage migration inhibitory factor [Brugia malayi] >gi_2190976 (AF002699) macrophage migration

inhibitory factor [Brugia malayi]

Seq. No.

8448 12774 1.R1040

Contig ID 5'-most EST

LIB3040-030-Q1-E1-H1

Seq. No.

8449

Contig ID

12784 1.R1040

5'-most EST

LIB3049-034-Q1-E1-E12

Seq. No.

8450

Contig ID 5'-most EST 12789 1.R1040 $leu70\overline{1}149449.h1$

Method

BLASTX

NCBI GI

g1718097 1090

BLAST score

E value

1.0e-119

Match length % identity

340 58

NCBI Description

VACUOLAR ATP SYNTHASE SUBUNIT AC39 (V-ATPASE AC39 SUBUNIT) (41 KD ACCESSORY PROTEIN) (DVA41) >gi_626048_pir_A55016

lysosomal membrane protein DVA41 - slime mold

(Dictyostelium discoideum) >gi_532733 (U13150) vacuolar

ATPase subunit DVA41 [Dictyostelium discoideum]



Contig ID 12793_1.R1040 5'-most EST bth700843731.h1

Method BLASTX
NCBI GI g3763916
BLAST score 377
E value 8.0e-36
Match length 260
% identity 33

NCBI Description (AC004450) unknown protein [Arabidopsis thaliana]

>gi 4531439 gb AAD22124.1 AC006224 6 (AC006224) unknown

protein [Arabidopsis thaliana]

Seq. No. 8452

Contig ID 12796_1.R1040 5'-most EST zzp700832237.h1

Method BLASTX
NCBI GI g4468813
BLAST score 387
E value 3.0e-37
Match length 202
% identity 39

NCBI Description (AL035601) putative protein [Arabidopsis thaliana]

Seq. No. 8453

Contig ID 12796 2.R1040

5'-most EST LIB3139-121-P1-N1-G4

Seq. No. 8454

Contig ID 12796_3.R1040 5'-most EST gsv701045344.h1

Seq. No. 8455

Contig ID 12802_1.R1040

5'-most EST LIB3040-030-Q1-E1-C5

Seq. No. 8456

Contig ID 12802_2.R1040 5'-most EST leu701145651.h1

Seq. No. 8457

Contig ID 12802_4.R1040

5'-most EST LIB3040-005-Q1-E1-A1

Seq. No. 8458

Contig ID 12808 1.R1040 5'-most EST asn701135489.h1

Seq. No. 8459

Contig ID 12810 1.R1040

5'-most EST LIB3170-012-Q1-K1-B10

Seq. No. 8460

Contig ID 12815_1.R1040 5'-most EST ncj700977424.h1

Method BLASTX NCBI GI g2731377

```
1
BLAST score
                   1.0e-21
E value
                   143
Match length
                   38
% identity
                  (U28739) similar to alcohol dehydrogenase/ribitol
NCBI Description
                   dehydrogenase [Caenorhabditis elegans]
                   8461
Seq. No.
                   12817 1.R1040
Contig ID
                   LIB3040-030-Q1-E1-A3
5'-most EST
                   BLASTN
Method
                   q2695738
NCBI GI
                   56
BLAST score
                   9.0e-23
E value
                   68
Match length
                   96
% identity
                   Pisum sativum mitochondrial tRNA-Gly gene and flanking
NCBI Description
                   sequences
Seq. No.
                   8462
                   12819 1.R1040
Contig ID
5'-most EST
                   ncj700975520.h1
                   8463
Seq. No.
                   12820_1.R1040
Contig ID
                   LIB3170-012-Q1-K1-B4
5'-most EST
Seq. No.
                   12831 1.R1040
Contig ID
                   uC-gmrominsoy316e04b1
5'-most EST
                   BLASTX
Method
                   g4544409
NCBI GI
                   1156
 BLAST score
                   1.0e-127
 E value
                   376
 Match length
                    60
 % identity
                    (AC006955) putative transcription factor [Arabidopsis
 NCBI Description
                    thaliana]
                    8465
 Seq. No.
                    12831 2.R1040
 Contig ID
                    uC-gmrominsoy180c03b1
 5'-most EST
                    BLASTX
 Method
                    g4544409
 NCBI GI
                    194
 BLAST score
                    2.0e-14
 E value
                    69
 Match length
                    59
 % identity
                    (AC006955) putative transcription factor [Arabidopsis
 NCBI Description
                    thaliana]
```

0 37

Seq. No. 8466

Contig ID 125

12831_3.R1040 LIB3051-072-Q1-K1-E10

5'-most EST Method

BLASTX g2244917 285

NCBI GI BLAST score

1534



E value 5.0e-25 Match length 139 % identity 45

NCBI Description (297339) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8467

Contig ID 12831 4.R1040

5'-most EST LIB3093-034-Q1-K1-D7

Seq. No. 8468

Contig ID 12831_5.R1040 5'-most EST uC-gmropic042d01b1

Seq. No. 8469

Contig ID 12831 6.R1040

5'-most EST LIB30 $\overline{9}2$ -008-Q1-K1-F1

Method BLASTX
NCBI GI g2244917
BLAST score 161
E value 1.0e-10
Match length 65
% identity 46

NCBI Description (Z97339) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8470

Contig ID 12831_7.R1040 5'-most EST jex700905749.h1

Seq. No. 8471

Contig ID 12840 1.R1040

5'-most EST LIB3109-029-Q1-K1-G1

Method BLASTX
NCBI GI g3122357
BLAST score 370
E value 2.0e-35
Match length 98

% identity 70

NCBI Description PUTATIVE LIPOATE-PROTEIN LIGASE B (LIPOATE BIOSYNTHESIS

PROTEIN B) >gi_2494127 (AC002376) Contains similarity to

Mycobacterium LIPB gene (gb_Q104041). [Arabidopsis

thaliana]

Seq. No. 8472

Contig ID 12856_1.R1040

5'-most EST LIB3106-069-P1-K1-F8

Seq. No. 8473

Contig ID 12867 1.R1040

5'-most EST LIB3051-006-Q1-K1-A11

Seq. No. 8474

Contig ID 12872_1.R1040

5'-most EST LIB3093-016-Q1-K1-F11

Method BLASTX
NCBI GI g1665817
BLAST score 200
E value 2.0e-15



Match length 128 % identity 36

NCBI Description (D87466) Similar to S.cerevisiae hypothetical protein L3111

(S59316) [Homo sapiens]

Seq. No. 8475

Contig ID 12873_1.R1040

5'-most EST LIB3170-010-Q1-K1-A6

Seq. No. 8476

Contig ID 12873_2.R1040

5'-most EST LIB3107-006-Q1-K1-C10

Seq. No. 8477

Contig ID 12875_1.R1040

5'-most EST LIB3170-011-Q1-K1-A5

Seq. No. 8478

Contig ID 12882_1.R1040

5'-most EST LIB3040-029-Q1-E1-B3

Seq. No. 8479

Contig ID 12883_1.R1040

5'-most EST LIB3040-020-Q1-E1-B11

Method BLASTX
NCBI GI g4559346
BLAST score 407
E value 1.0e-39
Match length 122
% identity 61

NCBI Description (AC006585) early nodulin 16 [Arabidopsis thaliana]

Seq. No. 8480

Contig ID 12883_2.R1040 5'-most EST hrw701062307.h1

Seq. No. 8481

Contig ID 12884_1.R1040 5'-most EST kmv700741435.h1

Seq. No. 8482

Contig ID 12898_1.R1040

5'-most EST LIB3170-011-Q1-J1-A4

Seq. No. 8483

Contig ID 12905_1.R1040 5'-most EST uC-gmropic016c07b1

Method BLASTX
NCBI GI g3355468
BLAST score 526
E value 2.0e-53
Match length 123
% identity 88

NCBI Description (AC004218) putative ribosomal protein L35 [Arabidopsis

thaliana]



```
Contig ID
                   12905 2.R1040
5'-most EST
                   zhf700953832.h1
Method
                   BLASTX
                   q3355468
NCBI GI
BLAST score
                  179
E value
                   3.0e-13
Match length
                   47
                  79
% identity
                   (ACO04218) putative ribosomal protein L35 [Arabidopsis
NCBI Description
                  thalianal
Seq. No.
                   8485
                  12908 1.R1040
Contig ID
                  LIB3040-028-Q1-E1-E5
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2224899
BLAST score
                  277
                   3.0e-24
E value
                  186
Match length
% identity
                   41
                  (U67133) DNA-binding protein PcMYB1 [Petroselinum crispum]
NCBI Description
Seq. No.
                   8486
Contig ID
                   12913 1.R1040
5'-most EST
                  LIB3040-005-Q1-E1-G5
Method
                  BLASTX
NCBI GI
                   g1421741
BLAST score
                   377
                   2.0e-36
E value
Match length
                  77
% identity
                   87
NCBI Description
                  (U54770) cytochrome P450 homolog [Lycopersicon esculentum]
Seq. No.
                   8487
                   12916 1.R1040
Contig ID
5'-most EST
                  LIB3040-028-Q1-E1-F7
Method
                  BLASTX
NCBI GI
                   q3360289
BLAST score
                   155
E value
                   2.0e-15
Match length
                   77
% identity
                   62
NCBI Description
                   (AF023164) leucine-rich repeat transmembrane protein kinase
                   1 [Zea mays]
                   8488
Seq. No.
                   12919 1.R1040
Contig ID
5'-most EST
                  LIB3170-005-Q1-K1-H6
Method
                  BLASTX
NCBI GI
                   q2920587
BLAST score
                   1416
E value
                   1.0e-157
Match length
                   539
```

NCBI Description (AF038362) TBP-associated factor 172 [Homo sapiens] >gi 2995136 emb CAA04475 (AJ001017) TAFII170 [Homo

sapiens]

54

% identity



12920 1.R1040 Contig ID

LIB3040-008-Q1-E1-B4 5'-most EST

Method BLASTX NCBI GI g2443886 BLAST score 314 E value 1.0e-28 87 Match length 76

(AC002294) Unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

% identity

12920 2.R1040 Contig ID $2hf70\overline{0}961575.h1$ 5'-most EST

8490

Seq. No.

8491 Contig ID 12920 5.R1040 5'-most EST zhf700964127.hl

Seq. No.

Contig ID 12943 1.R1040

5'-most EST LIB3040-026-Q1-E1-D12

8492

Method BLASTX NCBI GI g2832625 BLAST score 759 9.0e-81 E value Match length 221 % identity 66

NCBI Description (AL021711) putative protein [Arabidopsis thaliana]

8493 Seq. No.

12959 1.R1040 Contig ID

5'-most EST LIB3040-027-Q1-E1-G10

Seq. No. 8494

12965 1.R1040 Contig ID

LIB3040-027-Q1-E1-G6 5'-most EST

Method BLASTX NCBI GI g2326363 BLAST score 301 E value 7.0e-36 Match length 121 65 % identity

(AJ001037) DNA-directed RNA polymerase [Arabidopsis NCBI Description

thaliana]

Seq. No. 8495

Contig ID 12970 1.R1040

5'-most EST LIB3040-027-Q1-E1-H2

8496 Seq. No.

12976 1.R1040 Contig ID $uaw70\overline{0}664514.h1$ 5'-most EST

BLASTN Method g2760166 NCBI GI

BLAST score 40



5.0e-13 E value Match length 131 % identity 87

Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: NCBI Description

MBK20, complete sequence [Arabidopsis thaliana]

Seq. No. Contig ID 8497 12977 1.R1040

5'-most EST

LIB3040-027-Q1-E1-E4

Method NCBI GI BLASTN g2196463

BLAST score

213

E value Match length 1.0e-116

% identity

537 86

NCBI Description Arabidopsis thaliana chloroplast trnC, rpoB & rpoCl genes

Seq. No.

8498

Contig ID

12987 1.R1040

5'-most EST

LIB3051-009-Q1-E1-F7

Seq. No.

8499

Contig ID 5'-most EST 12988 1.R1040

epx701110320.h1

Seq. No.

8500

Contig ID

12992 1.R1040

5'-most EST

 $LIB31\overline{0}6-025-Q1-K1-D7$

Method

BLASTX

NCBI GI

g1174870

BLAST score

268

E value

2.0e-23

Match length

% identity

68 75

NCBI Description

UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 8.0 KD PROTEIN

>gi_633685_emb_CAA55861_ (X79274) ubiquinol--cytochrome c

reductase [Solanum tuberosum]

Seq. No.

8501

Contig ID

12992 2.R1040

5'-most EST Method

LIB3053-002-Q1-B1-D1

NCBI GI

BLASTX g1174870

BLAST score

231

E value

3.0e-19

Match length

67 69

% identity NCBI Description

UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 8.0 KD PROTEIN

>qi 633685 emb CAA55861 (X79274) ubiquinol--cytochrome c

reductase [Solanum tuberosum]

Seq. No.

8502

Contig ID

12992 3.R1040

5'-most EST

jC-gmro02800040c03a1

Seq. No.

8503

Contig ID

12997_1.R1040



```
5'-most EST
                  LIB3040-027-Q1-E1-D5
Seq. No.
                  8504
                  13023 1.R1040
Contig ID
5'-most EST
                  g4283734
Method
                  BLASTX
NCBI GI
                  g881625
BLAST score
                  575
E value
                  5.0e-59
Match length
                  125
                  83
% identity
NCBI Description
                  (U29432) 22.8 kDa protein [Ipomoea trifida]
                  8505
Seq. No.
                  13027 1.R1040
Contig ID
5'-most EST
                  gsv701044414.h1
Method
                  BLASTX
NCBI GI
                  g3292831
BLAST score
                  604
E value
                  2.0e-85
Match length
                  250
% identity
                  59
                  (AL031018) putative serine/threonine kinase [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  8506
                  13027 2.R1040
Contig ID
                  jC-gmle01810087g01a1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2253010
BLAST score
                  542
E value
                  1.0e-55
Match length
                  144
% identity
                  68
NCBI Description
                   (Y14199) MAP3K delta-1 protein kinase [Arabidopsis
                  thaliana]
Seq. No.
                  8507
                  13030 1.R1040
Contig ID
5'-most EST
                  LIB3106-115-Q1-K1-E4
Method
                  BLASTX
NCBI GI
                  g2997684
BLAST score
                  270
E value
                  1.0e-23
Match length
                  67
                  72
% identity
NCBI Description
                   (AF053302) putative transcriptional co-activator
                   [Arabidopsis thaliana]
```

Contig ID 13032_1.R1040

5'-most EST LIB3170-007-Q1-K1-G11

Method BLASTX
NCBI GI g4388832
BLAST score 255
E value 6.0e-22
Match length 66



% identity 91
NCBI Description (AC006528) putative DNA replication licensing factor with an MCM family domain (prosite:PDOC00662) [Arabidopsis thaliana]

Seq. No. 8509

Contig ID 13033 1.R1040

5'-most EST LIB3093-037-Q1-K1-F12

Method BLASTX
NCBI GI g2826786
BLAST score 301
E value 4.0e-27
Match length 123
% identity 54

NCBI Description (Y10905) RAPB protein [Oryza sativa]

Seq. No. 8510

Contig ID 13033_3.R1040 5'-most EST bth700849432.h1

Seq. No. 8511

Contig ID 13034 1.R1040

5'-most EST LIB3170-007-Q1-K1-G12

Seq. No. 8512

Contig ID 13036 1.R1040

5'-most EST LIB3170-005-Q1-K1-G3

Method BLASTX
NCBI GI g3582021
BLAST score 473
E value 2.0e-47
Match length 149
% identity 60

NCBI Description (Y09423) cytochrome P450 [Nepeta racemosa]

Seq. No. 8513

Contig ID 13041 1.R1040

5'-most EST LIB3040-026-Q1-E1-D2

Method BLASTX
NCBI GI g3935184
BLAST score 194
E value 8.0e-15
Match length 93
% identity 49

NCBI Description (AC004557) F17L21.27 [Arabidopsis thaliana]

Seq. No. 8514

Contig ID 13045 1.R1040

5'-most EST jC-qmle01810094d11a1

Method BLASTX
NCBI GI g3021285
BLAST score 250
E value 2.0e-21
Match length 75
% identity 65

NCBI Description (AL022347) hypothetical protein [Arabidopsis thaliana]

>gi_3451058_emb_CAA20454.1_ (AL031326) hypothetical protein



[Arabidopsis thaliana]

8515 Seq. No.

13054 1.R1040 Contig ID

LIB3040-026-Q1-E1-C9 5'-most EST

8516 Seq. No.

13060 1.R1040 Contig ID

iC-qmf102220141d09a1 5'-most EST

8517 Seq. No.

13061 1.R1040 Contig ID

LIB3040-026-Q1-E1-B6 5'-most EST

8518 Seq. No.

13063 1.R1040 Contig ID

uC-gmrominsoy299a01b1 5'-most EST

8519 Seq. No.

13075 1.R1040 Contig ID

LIB3040-025-Q1-E1-H12 5'-most EST

8520 Seq. No.

13079 1.R1040 Contig ID

LIB3093-036-Q1-K1-A2 5'-most EST

BLASTN Method g11772 NCBI GI 77 BLAST score 9.0e-35 E value 365 Match length

% identity

NCBI Description Tobacco chloroplast tRNA-Ser and tRNA-Gln genes

8521 Seq. No.

13079 2.R1040 Contig ID

LIB3170-071-Q1-K1-E12 5'-most EST

BLASTN Method NCBI GI g296156 BLAST score 77 4.0e-35 E value 204 Match length 85 % identity

Spinach chloroplast psbK gene for low-molecular mass NCBI Description

protein K of photosystem II

8522 Seq. No.

13084_1.R1040 Contig ID

LIB3040-026-Q1-E1-A3 5'-most EST

BLASTX Method g3257978 NCBI GI 511 BLAST score 2.0e-51 E value 230 Match length 46 % identity

(AP000006) 249aa long hypothetical protein [Pyrococcus NCBI Description

horikoshii]



Contig ID 13084 2.R1040 5'-most EST leu701151015.h1

Method BLASTX
NCBI GI g2650133
BLAST score 153
E value 6.0e-10
Match length 55
% identity 56

NCBI Description (AE001070) ribonuclease PH (rph) [Archaeoglobus fulgidus]

Seq. No. 8524

Contig ID 13090 1.R1040

5'-most EST LIB3040-025-Q1-E1-F10

Seq. No. 8525

Contig ID 13095 1.R1040

5'-most EST LIB3040-025-Q1-E1-F9

Method BLASTX
NCBI GI g2832625
BLAST score 156
E value 1.0e-21
Match length 92
% identity 64

NCBI Description (AL021711) putative protein [Arabidopsis thaliana]

Seq. No. 8526

Contig ID 13096_1.R1040 5'-most EST jsh701063926.h1

Method BLASTX
NCBI GI g3647283
BLAST score 561
E value 1.0e-57
Match length 135
% identity 79

NCBI Description (AJ011418) ubiquitin activating enzyme [Lycopersicon

esculentum]

Seq. No. 8527

Contig ID 13096_2.R1040 5'-most EST bth700848190.h1

Method BLASTX
NCBI GI 94455237
BLAST score 359
E value 4.0e-34
Match length 144
% identity 55

NCBI Description (AL035523) ubiquitin activating enzyme-like protein

[Arabidopsis thaliana]

Seq. No. 8528

Contig ID 13096_3.R1040

5'-most EST LIB3049-034-Q1-E1-H1

Method BLASTX
NCBI GI g4455237
BLAST score 355
E value 9.0e-34



Match length 97 % identity 71

NCBI Description (AL035523) ubiquitin activating enzyme-like protein

[Arabidopsis thaliana]

Seq. No. 8529

Contig ID 13098 1.R1040

5'-most EST LIB3106-076-Q1-K1-C3

Method BLASTX
NCBI GI g548774
BLAST score 1027
E value 1.0e-112
Match length 240
% identity 83

NCBI Description 60S RIBOSOMAL PROTEIN L7A >gi_542158_pir__S38360 ribosomal

protein L7a - rice >gi_303855_dbj_BAA02156_ (D12631)

ribosomal protein L7A [Oryza sativa]

Seq. No. 8530

Contig ID 13100 1.R1040

5'-most EST LIB3170-005-Q1-J1-B2

Seq. No. 8531

Contig ID 13104 1.R1040

5'-most EST LIB3170-006-Q1-K1-F4

Seq. No. 8532

Contig ID 13105_1.R1040

5'-most EST LIB3040-025-Q1-E1-C8

Seq. No. 8533

Contig ID 13107_1.R1040

NCBI Description (AC004218) putative ribosomal protein L35 [Arabidopsis

thaliana]

Seq. No. 8534

Contig ID 13107_2.R1040
5'-most EST 94396498
Method BLASTX
NCBI GI 93355468
BLAST score 515
E value 4.0e-52
Match length 122

% identity 87

NCBI Description (AC004218) putative ribosomal protein L35 [Arabidopsis

thaliana]

Seq. No. 8535

Contig ID 13110 1.R1040

5'-most EST LIB3107-003-Q1-K1-B4



Method BLASTX NCBI GI g464981 BLAST score 752 5.0e-80 E value 147 Match length % identity 93

UBIQUITIN-CONJUGATING ENZYME E2-17 KD (UBIQUITIN-PROTEIN NCBI Description LIGASE) (UBIQUITIN CARRIER PROTEIN) >gi 388207 (L23762)

ubiquitin carrier protein [Lycopersicon esculentum]

Seq. No. 8536

Contig ID 13115 1.R1040 5'-most EST leu701157895.h1

Method BLASTX NCBI GI q4490308 BLAST score 1004 E value 1.0e-109 Match length 303 % identity

(AL035678) putative protein [Arabidopsis thaliana] NCBI Description

8537 Seq. No.

13116 1.R1040 Contig ID 5'-most EST bth700848141.h1

Method BLASTX g3687243 NCBI GI BLAST score 256 E value 7.0e-22 Match length 67 % identity 76

(AC005169) putative ribosomal protein [Arabidopsis NCBI Description

thaliana]

Seq. No. 8538

13116 2.R1040 Contig ID 5'-most EST pmv700888753.h1

Method BLASTX NCBI GI q3687243 BLAST score 254 9.0e-22 E value 66 Match length 74 % identity

(AC005169) putative ribosomal protein [Arabidopsis NCBI Description

thaliana]

Seq. No. 8539

Contig ID 13121 1.R1040

5'-most EST uC-gmflminsoy008b03b1

8540 Seq. No.

Contig ID 13124 1.R1040

5'-most EST LIB3170-007-Q1-K1-F9

8541 Seq. No.

13128 1.R1040 Contig ID

LIB3170-006-Q1-K1-F1 5'-most EST



8542

Contig ID 5'-most EST 13130 1.R1040 zsg701123209.h1

Method NCBI GI BLAST score BLASTX g2559012 375

E value Match length 1.0e-35 128

58

% identity NCBI Description

(AF026293) chaperonin containing t-complex polypeptide 1,

beta subunit; CCT-beta [Homo sapiens] >gi 4090929

(AF026166) chaperonin-containing TCP-1 beta subunit homolog

[Homo sapiens]

Seq. No. Contig ID 8543

13133 1.R1040

5'-most EST

LIB3170-007-Q1-K1-F12

Seq. No. Contig ID 8544

5'-most EST

13134 1.R1040 $leu70\overline{1}144349.h1$

Method BLASTN NCBI GI q4388705 BLAST score 37 4.0e-11 E value Match length 142

% identity

89

NCBI Description

Arabidopsis thaliana chromosome I BAC F20D21 genomic sequence, complete sequence [Arabidopsis thaliana]

Seq. No. Contig ID 8545

13134 3.R1040

5'-most EST

LIB3138-013-Q1-N1-G2

Method BLASTX NCBI GI g4115731 BLAST score 160 7.0e-11 E value 86 Match length % identity 44

NCBI Description

(AB017507) Apg12 [Homo sapiens]

Seq. No.

8546

Contig ID

13137 1.R1040

5'-most EST

LIB3040-025-Q1-E1-C11

Seq. No.

8547

Contig ID 5'-most EST 13137 2.R1040 epx701105423.hl

Seq. No.

8548

Contig ID

13140 1.R1040

5'-most EST

LIB3092-053-Q1-K1-D9

Seq. No.

8549

Contig ID

13147 1.R1040

5'-most EST

g5676941



Contig ID 13147 2.R1040

5'-most EST LIB3074-025-Q1-E1-D1

Method BLASTN
NCBI GI g728715
BLAST score 75
E value 7.0e-34
Match length 191

NCBI Description S.oleracea mRNA for 6.1 kDa polypeptide of photosystem II

Seq. No. 855

% identity

Contig ID 13147 4.R1040

5'-most EST LIB3074-028-Q1-K1-A9

85

Method BLASTX
NCBI GI g2493694
BLAST score 180
E value 4.0e-13
Match length 98

% identity 42

NCBI Description PHOTOSYSTEM II REACTION CENTRE W PROTEIN PRECURSOR (PSII

6.1 KD PROTEIN) >gi_1076268_pir__S53025 photosystem II protein - spinach >gi_728716_emb_CAA59409_ (X85038) protein

of photosystem II [Spinacia oleracea]

Seq. No. 8552

Contig ID 13149 1.R1040

5'-most EST LIB31 $\overline{7}$ 0-007-Q1-K1-F3

Method BLASTX
NCBI GI g3328587
BLAST score 334
E value 1.0e-33
Match length 165
% identity 50

NCBI Description (AE001292) CMP-2-keto-3-deoxyoctulosonic acid synthetase

[Chlamydia trachomatis]

Seq. No. 8553

Contig ID 13151_1.R1040

5'-most EST LIB3040-024-Q1-E1-G5

Seq. No. 8554

Contig ID 13158_1.R1040

5'-most EST LIB3138-015-Q1-N2-E2

Seq. No. 8555

Contig ID 13158 2.R1040

5'-most EST uC-gmflminsoy043f05b1

Seq. No. 8556

Contig ID 13158_3.R1040 5'-most EST leu701154395.h1

Seq. No. 8557

Contig ID 13159_1.R1040 5'-most EST uC-gmropic041a09b1



Seq. No. Contig ID

8558

13159 2.R1040 vwf700677224.h1

Seq. No. Contig ID

5'-most EST

8559

13160 1.R1040 $jex70\overline{0}906127.h1$ 5'-most EST

BLASTX Method NCBI GI g3402690 BLAST score 233 E value 4.0e-19 Match length 94 % identity

NCBI Description

(AC004697) hypothetical protein, 3' partial [Arabidopsis

thaliana]

Seq. No. Contig ID 8560

5'-most EST

13161 1.R1040 zzp700832222.h1

Method BLASTX NCBI GI g1352663 BLAST score 183 3.0e-13E value Match length 48 % identity 73

SERINE/THREONINE PROTEIN PHOSPHATASE PP2A-3 CATALYTIC NCBI Description

SUBUNIT >gi_1076388_pir__S52659 phosphoprotein phosphatase

(EC 3.1.3.16) 2A isoform 3 - Arabidopsis thaliana >gi 466441 (M96841) Ser/Thr protein phosphatase

[Arabidopsis thaliana]

>qi 4559341 gb AAD23003.1 AC007087 22 (AC007087) serine/threonine protein phosphatase PP2A-3 catalytic

subunit [Arabidopsis thaliana]

>gi_4567320_gb_AAD23731.1_AC005956 20 (AC005956)

serine/threonine protein phosphatase [Arabidopsis thaliana]

Seq. No.

8561

13161 2.R1040 Contig ID

5'-most EST uC-qmrominsoy136h01b1 Method BLASTX NCBI GI g1352664 BLAST score 776 E value 1.0e-82 160

Match length % identity 91

SERINE/THREONINE PROTEIN PHOSPHATASE PP2A-4 CATALYTIC NCBI Description

SUBUNIT >gi 2117984 pir S52660 phosphoprotein phosphatase

(EC 3.1.3.16) 2A isoform 4 - Arabidopsis thaliana >qi 473259 (U08047) Ser/Thr protein phosphatase [Arabidopsis thaliana] >gi 4204949 (U60136)

serine/threonine protein phosphatase 2A-4 catalytic subunit

[Arabidopsis thaliana]

Seq. No.

8562

Contig ID 5'-most EST 13169 1.R1040 kmv700737865.h1



Contig ID 13173 1.R1040

5'-most EST LIB3051-027-Q1-K1-G8

Method BLASTX
NCBI GI g4056425
BLAST score 744
E value 9.0e-79
Match length 164
% identity 82

NCBI Description (AC005322) ESTs gb_H36249, gb_AA59732 and gb_AA651219 come

from this gene. [Arabidopsis thaliana]

Seq. No. 8564

Contig ID 13173 2.R1040

5'-most EST LIB3170-006-Q1-K1-E8

Seq. No. 8565

Contig ID 13178 1.R1040

5'-most EST LIB3072-051-Q1-E1-C1

Method BLASTN
NCBI GI g2764803
BLAST score 146
E value 2.0e-76
Match length 366
% identity 85

NCBI Description G.max mRNA for epoxide hydrolase

Seq. No. 8566

Contig ID 13182 1.R1040

5'-most EST LIB3040-023-Q1-E1-G8

Method BLASTX
NCBI GI g3702329
BLAST score 156
E value 2.0e-10

Match length 50 % identity 62

NCBI Description (AC005397) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8567

Contig ID 13183_1.R1040 5'-most EST hrw701056967.h1

Method BLASTX
NCBI GI g2811026
BLAST score 357
E value 9.0e-34
Match length 80
% identity 88

NCBI Description TCP1-CHAPERONIN COFACTOR A HOMOLOG >gi 1946375 (U93215)

TCP1-chaperonin cofactor A isolog [Arabidopsis thaliana] >gi_2347204 (AC002338) TCP1-chaperonin cofactor A isolog

[Arabidopsis thaliana]

Seq. No. 8568

Contig ID 13183 2.R1040

5'-most EST LIB3040-023-Q1-E1-G9

Method BLASTX NCBI GI g2811026



BLAST score 353 1.0e-33 E value Match length 78

% identity 88

NCBI Description

TCP1-CHAPERONIN COFACTOR A HOMOLOG >gi 1946375 (U93215)

TCP1-chaperonin cofactor A isolog [Arabidopsis thaliana] >gi 2347204 (AC002338) TCP1-chaperonin cofactor A isolog

[Arabidopsis thaliana]

Seq. No.

Contig ID

13200 1.R1040

5'-most EST

LIB3170-006-Q1-K1-C2 BLASTX

Method NCBI GI BLAST score

q4567248 888 1.0e-95

8569

E value Match length % identity

303 57

8570

NCBI Description

(AC007070) unknown protein [Arabidopsis thaliana]

Seq. No.

Contig ID

13214 1.R1040

5'-most EST

LIB3138-061-Q1-N1-C5

Seq. No.

8571

Contig ID

13214 2.R1040

5'-most EST

jC-gmle01810027g04d1

Seq. No.

8572

Contig ID

13214 3.R1040

5'-most EST

LIB3170-005-Q1-K1-D8

Seq. No.

8573

Contig ID 5'-most EST 13214 4.R1040 jC-gmle01810061e02a1

Seq. No.

8574

Contig ID 5'-most EST 13216 1.R1040 $uxk70\overline{0}672719.h1$

Seq. No.

8575

Contig ID 5'-most EST 13216 2.R1040 uxk700672738.h1

Seq. No.

8576

Contig ID 5'-most EST

13216 3.R1040 awf700840091,h1

Seq. No.

8577

13222 1.R1040

Contig ID 5'-most EST

LIB3107-002-Q1-K1-B3

Seq. No.

8578

Contig ID

13227 1.R1040

5'-most EST

LIB3170-006-Q1-J1-C12

Seq. No.

8579



Contig ID 13235 1.R1040

5'-most EST LIB3040-022-Q1-E1-H5

Method BLASTX
NCBI GI g2232057
BLAST score 232
E value 5.0e-19
Match length 69
% identity 67

NCBI Description (AF000177) CaSm [Homo sapiens]

Seq. No. 8580

Contig ID 13235_2.R1040 5'-most EST jex700905452.h1

Seq. No. 8581

Contig ID 13242 1.R1040

5'-most EST LIB3049-052-Q1-E1-G2

Seq. No. 8582

Contig ID 13244 1.R1040

5'-most EST LIB3170-004-Q1-K1-A10

Method BLASTX
NCBI GI g4490737
BLAST score 509
E value 2.0e-51
Match length 184
% identity 34

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 8583

Contig ID 13256 1.R1040

5'-most EST LIB3040-022-Q1-E1-G3

Seq. No. 8584

Contig ID 13260_1.R1040 5'-most EST epx701105025.h1

Seq. No. 8585

Contig ID 13266_1.R1040 5'-most EST wrg700786736.h2

Seq. No. 8586

Contig ID 13270 1.R1040

5'-most EST LIB3170-005-Q1-K1-H4

Seq. No. 8587

Contig ID 13274 1.R1040

5'-most EST LIB3170-006-Q1-J1-E6

Seq. No. 8588

Contig ID 13285 1.R1040

5'-most EST LIB3040-022-Q1-E1-A10

Seq. No. 8589

Contig ID 13288 1.R1040

5'-most EST LIB3040-022-Q1-E1-A6

Method BLASTX



NCBI GI g465602 BLAST score 261 E value 1.0e-22 Match length 98 % identity 57

NCBI Description HYPOTHETICAL 32.9 KD PROTEIN IN NFO-FRUA INTERGENIC REGION

>gi_405885 (U00007) yeiN [Escherichia coli] >gi_1788490
(AE000306) orf, hypothetical protein [Escherichia coli]
>gi_744200_prf__2014253BL yeiN gene [Escherichia coli]

Seq. No. 8590

Contig ID 13295_1.R1040 5'-most EST uC-gmronoir057d05b1

Seq. No. 8591

Contig ID 13295 2.R1040

5'-most EST LIB3051-025-Q1-K1-G11

Method BLASTX
NCBI GI g3786008
BLAST score 175
E value 2.0e-12
Match length 127
% identity 39

NCBI Description (AC005499) unknown protein [Arabidopsis thaliana]

Seq. No. 8592

Contig ID 13300 1.R1040

5'-most EST LIB3055-012-Q1-N1-B12

Method BLASTX
NCBI GI g3335060
BLAST score 609
E value 5.0e-63
Match length 225
% identity 53

NCBI Description (AF025842) plasma membrane-type calcium ATPase [Arabidopsis

thaliana] >gi_4468989_emb_CAB38303 (AL035605) plasma

membrane-type calcium ATPase (ACA2) [Arabidopsis thaliana]

Seq. No. 8593

Contig ID 13300 2.R1040

5'-most EST LIB3040-021-Q1-E1-F2

Method BLASTX
NCBI GI g3098571
BLAST score 519
E value 2.0e-52
Match length 214
% identity 47

NCBI Description (AF049028) BURP domain containing protein [Brassica napus]

Seq. No. 8594

Contig ID 13300 4.R1040 5'-most EST seb700651485.h1

Seq. No. 8595

Contig ID 13304_1.R1040 5'-most EST awf700840276.h1

Method BLASTX



NCBI GI g3132470
BLAST score 179
E value 5.0e-13
Match length 96
% identity 50

NCBI Description (AC003096) unknown protein [Arabidopsis thaliana]

Seq. No. 8596

Contig ID 13304 2.R1040

5'-most EST uC-gmflminsoy084b04b1

Seq. No. 8597

Contig ID 13308_1.R1040 5'-most EST hrw701060302.h1

Method BLASTX
NCBI GI g2739376
BLAST score 263
E value 8.0e-23
Match length 90
% identity 58

NCBI Description (AC002505) putative permease [Arabidopsis thaliana]

Seq. No. 8598

Contig ID 13316 1.R1040 5'-most EST kl1701208551.h1

Method BLASTX
NCBI GI g2132842
BLAST score 344
E value 4.0e-32
Match length 118
% identity 53

NCBI Description probable membrane protein YOL077c - yeast (Saccharomyces

cerevisiae) >gi_1419909_emb_CAA99087_ (Z74819) ORF YOL077c

[Saccharomyces cerevisiae]

Seq. No. 8599

Contig ID 13323_1.R1040 5'-most EST hrw701060978.h1

Seq. No. 8600

Contig ID 13324 1.R1040

5'-most EST LIB3040-021-Q1-E1-D4

Seq. No. 8601

Contig ID 13335_1.R1040

5'-most EST LIB3170-007-Q1-K1-C5

Method BLASTX
NCBI GI 9499693
BLAST score 491
E value 2.0e-49
Match length 167
% identity 65

NCBI Description (L32095) cyclophilin [Vicia faba]

Seq. No. 8602

Contig ID 13337 1.R1040

5'-most EST LIB3040-021-Q1-E1-A4



```
Seq. No.
                  8603
                  13339 1.R1040
Contig ID
5'-most EST
                  LIB3170-006-Q1-K1-C1
Method
                  BLASTX
NCBI GI
                  g2344897
BLAST score
                  150
E value
                  1.0e-09
                  100
Match length
                  43
% identity
NCBI Description (AC002388) unknown protein [Arabidopsis thaliana]
Seq. No.
                  8604
                  13342 1.R1040
Contig ID
                  rlr700897809.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4158232
BLAST score
                  548
                  4.0e-56
E value
Match length
                  110
% identity
                  88
                  (Y18626) reversibly glycosylated polypeptide [Triticum
NCBI Description
                  aestivum]
                  8605
Seq. No.
Contig ID
                  13345 1.R1040
5'-most EST
                  LIB3040-021-Q1-E1-C12
Method
                  BLASTX
NCBI GI
                  g2739371
                  293
BLAST score
                  2.0e-26
E value
                  68
Match length
                  84
% identity
NCBI Description
                  (AC002505) unknown protein [Arabidopsis thaliana]
                  8606
Seq. No.
                  13364 1.R1040
Contig ID
5'-most EST
                  uC-qmflminsoy036f09b1
Method
                  BLASTX
NCBI GI
                  g2244834
BLAST score
                  186
E value
                  1.0e-13
Match length
                  96
% identity
                  53
                  (Z97337) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  8607
Contig ID
                  13381 1.R1040
5'-most EST
                  LIB3040-020-Q1-E1-E1
                  8608
Seq. No.
Contig ID
                  13385 1.R1040
5'-most EST
                  LIB3040-020-Q1-E1-C6
```

Method BLASTX g4263701 NCBI GI BLAST score 335 3.0e-31 E value



Match length 110 % identity 64

(AC006223) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No.

8609 13388 1.R1040

Contig ID 5'-most EST

LIB3049-028-Q1-E1-G3

Method NCBI GI BLAST score BLASTX g3650030 255

1.0e-21

E value Match length % identity

78 59

NCBI Description

(AC005396) unknown protein [Arabidopsis thaliana]

Seq. No.

8610

Contig ID

13396 1.R1040

5'-most EST

LIB3040-020-Q1-E1-D8

Seq. No.

8611

Contig ID

13415 1.R1040

5'-most EST

LIB3040-020-Q1-E1-C1

Seq. No.

8612

Contig ID

13416 1.R1040

5'-most EST

LIB3040-019-Q1-E1-F7

Method NCBI GI BLASTX g4249662

BLAST score

275

E value

2.0e-46

Match length % identity

118 84

NCBI Description

(AF089810) Altered Response to Gravity [Arabidopsis

thaliana]

Seq. No.

8613

Contig ID

13421 1.R1040

5'-most EST

LIB3040-019-Q1-E1-G11

Seq. No.

8614

Contig ID

13422 1.R1040

5'-most EST

LIB3106-111-Q1-K1-F11

Method NCBI GI BLAST score BLASTX g132918

E value

334

Match length

5.0e - 3175

% identity NCBI Description 87 50S RIBOSOMAL PROTEIN L35, CHLOROPLAST PRECURSOR (CL35)

>gi_81486_pir__A36107 ribosomal protein L35 precursor, chloroplast - spinach >gi 170139 (M60449) ribosomal protein

L35 [Spinacia oleracea]

Seq. No.

8615

Contig ID

13422 2.R1040

5'-most EST

LIB3040-011-Q1-E1-B5

Method

BLASTX



g132918 NCBI GI BLAST score 314 1.0e-28 E value Match length 102 73 % identity

NCBI Description 50S RIBOSOMAL PROTEIN L35, CHLOROPLAST PRECURSOR (CL35) >gi_81486_pir__A36107 ribosomal protein L35 precursor,

chloroplast - spinach >gi 170139 (M60449) ribosomal protein

L35 [Spinacia oleracea]

Seq. No.

8616 13426 1.R1040 Contig ID

5'-most EST LIB3109-018-Q1-K1-A4

Seq. No.

13427 1.R1040 Contig ID

5'-most EST LIB3073-021-Q1-K1-C1

8617

Seq. No.

8618 Contig ID 13433 1.R1040 5'-most EST bth700848210.hl

Method BLASTX NCBI GI g2144099 BLAST score 188 3.0e-14 E value Match length 77 % identity 42

NCBI Description Set alpha isoform - rat >gi_545263_bbs_143660 (S68589) Set

alpha isoform=leukemogenesis protein {alternatively

spliced} [rats, neonatal kidney, Peptide, 289 aa] [Rattus sp.] >gi 741750 prf 2008109A set gene [Rattus norvegicus]

Seq. No.

Contig ID 13439 1.R1040

5'-most EST LIB3073-003-Q1-K1-F3

8619

Method BLASTX NCBI GI q585960 BLAST score 206 E value 4.0e-16 Match length 42 % identity 93

PROTEIN TRANSPORT PROTEIN SEC61 BETA SUBUNIT NCBI Description

>gi 433665 emb CAA81412 (Z26753) Sec61 beta-subunit

homolog [Arabidopsis thaliana]

8620 Seq. No.

13442 1.R1040 Contig ID zzp700836203.hl 5'-most EST

Method BLASTX NCBI GI g3885334 BLAST score 182 E value 2.0e-13 Match length 48 % identity 65

NCBI Description (AC005623) putative argonaute protein [Arabidopsis

thaliana]



```
Seq. No.
```

8621

13446 1.R1040 Contig ID

5'-most EST

LIB3040-019-Q1-E1-F12

Seq. No.

8622

Contig ID 5'-most EST 13452 1.R1040

Method

LIB3040-019-Q1-E1-B8 BLASTX

NCBI GI

q1914685

BLAST score E value

352 2.0e-33

Match length % identity

74 89

NCBI Description

(Y12014) RAD23 protein, isoform II [Daucus carota]

Seq. No.

8623

Contig ID

13456 1.R1040

5'-most EST

LIB3170-007-Q1-K1-B6

Seq. No.

8624

Contig ID

13467 1.R1040

5'-most EST

LIB3040-018-Q1-E1-H11

Seq. No.

8625

Contig ID

13475 1.R1040

5'-most EST

LIB3065-003-Q1-N1-G12

Seq. No.

8626

Contig ID 5'-most EST 13476 1.R1040 zhf700963568.h1

Method

BLASTX

NCBI GI

g4101589

BLAST score

571

E value

1.0e-58

Match length

205

% identity

52

NCBI Description

(AF005050) aspartyl aminopeptidase [Homo sapiens]

Seq. No.

8627

Contig ID 5'-most EST 13479 2.R1040 asn701133559.h2

Method

BLASTX

NCBI GI

g132964

BLAST score

212

E value

9.0e-17

Match length

52

% identity

73

NCBI Description

60S RIBOSOMAL PROTEIN L29 (P23) >gi 71376 pir R6RT43

ribosomal protein RL43 - rat >gi 57145 emb CAA43146

(X60744) ribosomal protein [Rattus norvegicus]

>gi_312208_emb_CAA48344 (X68283) rat ribosomal protein L29

[Rattus norvegicus]

Seq. No.

8628

Contig ID

13488 1.R1040

5'-most EST

LIB3040-018-Q1-E1-F3



Contig ID 13489 1.R1040

5'-most EST LIB3051-117-Q1-K1-F11

Method BLASTX g1705678 NCBI GI BLAST score 1079 E value 1.0e-161 Match length 312 72 % identity

CELL DIVISION CYCLE PROTEIN 48 HOMOLOG (VALOSIN CONTAINING NCBI Description

> PROTEIN HOMOLOG) (VCP) >gi_862480 (U20213) valosin-containing protein [Glycine max]

Seq. No.

13490 1.R1040 Contig ID 5'-most EST gsv701054670.hl

Seq. No. 8631

13495 1.R1040 Contig ID

5'-most EST LIB3040-018-Q1-E1-E5

Method BLASTX NCBI GI q3805765 BLAST score 224 3.0e-18 E value Match length 54 % identity 76

NCBI Description (AC005693) putative protein kinase [Arabidopsis thaliana]

Seq. No. 8632

13498 1.R1040 Contig ID 5'-most EST smc700750158.hl

Method BLASTX NCBI GI g2501538 BLAST score 153 E value 6.0e-10 Match length 91 % identity 40

HYPOTHETICAL 51.0 KD PROTEIN SLL0996 NCBI Description

>gi 1652230 dbj BAA17153 (D90904) hypothetical protein

[Synechocystis sp.]

Seq. No. 8633

13500 1.R1040 Contig ID

5'-most EST LIB3040-018-Q1-E1-C4

Seq. No. 8634

Contig ID

13501 1.R1040 LIB3040-018-Q1-E1-C6 5'-most EST

BLASTX Method NCBI GI g3122853 BLAST score 383 1.0e-36 E value Match length 250 25 % identity

NCBI Description WD-REPEAT PROTEIN SAZD >gi 1082863 pir A49367 transducin

homolog sazD - human >gi_414536 (U02609) transducin-like

protein [Homo sapiens]



Contig ID 13512 1.R1040

5'-most EST LIB3170-005-Q1-J1-B3

Seq. No. 8636

Contig ID 13514 1.R1040

5'-most EST LIB3049-051-Q1-E1-A9

Method BLASTX
NCBI GI g2135057
BLAST score 416
E value 2.0e-40
Match length 226
% identity 41

NCBI Description EB1 - human >gi 998357 (U24166) EB1 [Homo sapiens]

Seq. No.

Contig ID 13520_1.R1040

5'-most EST LIB3170-006-Q1-J1-A1

8637

Seq. No. 8638

Contig ID 13524 1.R1040

5'-most EST LIB3139-075-P1-N1-G4

Seq. No. 8639

Contig ID 13524 2.R1040

5'-most EST LIB3051-117-Q1-K1-B9

Seq. No. 8640

Contig ID 13525 1.R1040

5'-most EST LIB3040-018-Q1-E1-B2

Seq. No. 8641

Contig ID 13532_1.R1040

5'-most EST LIB3040-018-Q1-E1-A2

Seq. No. 8642

Contig ID 13534_1.R1040

Seq. No. 8643

Contig ID 13534 2.R1040

5'-most EST LIB3040-017-Q1-E1-G10

Seq. No. 8644

Contig ID 13534 3.R1040

5'-most EST LIB3092-044-Q1-K1-G11

Seq. No. 8645

Contig ID 13538 1.R1040 5'-most EST trc700566156.h1

Seq. No. 8646

Contig ID 13547_1.R1040 5'-most EST g5677607 Method BLASTX NCBI GI g3935148



BLAST score 786 6.0e-84E value Match length 228 73 % identity

NCBI Description (AC005106) T25N20.12 [Arabidopsis thaliana]

Seq. No.

8647

Contig ID

13551_1.R1040

5'-most EST

LIB3170-061-Q1-J1-E1

Method NCBI GI BLAST score BLASTN q3599418

E value Match length

1.0e-26 80

% identity

95

63

NCBI Description

Glycine max alternative oxidase precursor (Aox1) gene,

nuclear gene encoding mitochondrial protein, complete cds

şě.

Seq. No.

8648

Contig ID 5'-most EST 13558 1.R1040 zsg701124685.h1

Method BLASTX NCBI GI q4580466 BLAST score 372 E value 2.0e-35 Match length 100 % identity 74

(AC006081) putative 50S ribosomal protein L4 [Arabidopsis NCBI Description

thaliana]

Seq. No.

8649

Contig ID 5'-most EST 13558 2.R1040

LIB3040-017-Q1-E1-F2

Seq. No.

8650

13575 1.R1040 Contig ID

5'-most EST

LIB3040-017-Q1-E1-D2

Seq. No.

8651

13577 1.R1040 Contig ID

5'-most EST

LIB3040-017-Q1-E1-D7

Seq. No.

8652

Contig ID

13579 1.R1040

5'-most EST

LIB3040-017-Q1-E1-C12

Method BLASTX NCBI GI g3201610 BLAST score 323 5.0e-30 E value 70 Match length % identity 77

NCBI Description

(AC004669) unknown protein [Arabidopsis thaliana]

Seq. No.

8653

Contig ID

13583 1.R1040

5'-most EST

LIB3170-002-Q1-K1-H8

Method

BLASTX

4.

```
q2565436
NCBI GI
BLAST score
                  423
E value
                  2.0e-41
                  91
Match length
                  88
% identity
                  (AF028842) DegP protease precursor [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  8654
                  13608 1.R1040
Contig ID
5'-most EST
                  asn701135465.h1
                  BLASTX
Method
NCBI GI
                  g2213595
BLAST score
                  481
                  5.0e-48
E value
Match length
                  219
% identity
                  44
NCBI Description
                  (AC000348) T7N9.15 [Arabidopsis thaliana]
Seq. No.
                  8655
```

13608 2.R1040 Contig ID

5'-most EST bth700849093.h1 Method BLASTX NCBI GI g2213595 BLAST score 244 E value 4.0e-28 Match length 112 % identity 58

NCBI Description (AC000348) T7N9.15 [Arabidopsis thaliana]

8656 Seq. No. 13608 3.R1040 Contig ID 5'-most EST pxt700944327.h1 Method BLASTX NCBI GI q2190992 BLAST score 220 E value 3.0e-18

Match length 65 69 % identity

(AF004358) glutathione S-transferase TSI-1 [Aegilops NCBI Description tauschii]

Seq. No. 8657

Contig ID 13609 1.R1040 5'-most EST rlr700895595.h1

Method BLASTX NCBI GI q4006860 BLAST score 666 8.0e-70 E value Match length 184 % identity 69

NCBI Description (Z99707) thiol-disulfide interchange like protein

[Arabidopsis thaliana]

8658 Seq. No.

Contig ID 13616 1.R1040

5'-most EST LIB3170-003-Q1-J1-H12



```
Seq. No.
                  8659
                  13620 1.R1040
Contig ID
                  LIB3073-019-Q1-K1-A5
5'-most EST
Seq. No.
                  8660
                  13620 2.R1040
Contig ID
                  zvp700764268.hl
5'-most EST
                  8661
Seq. No.
                  13622 1.R1040
Contig ID
                  LIB3040-016-Q1-E1-F4
5'-most EST
                  BLASTX
Method
                  g542070
NCBI GI
                   154
BLAST score
                   4.0e-10
E value
Match length
                   56
                   54
% identity
NCBI Description cytochrome P450 77A1 - eggplant
                   8662
Seq. No.
                   13625 1.R1040
Contiq ID
                   LIB3170-001-Q1-J1-H2
5'-most EST
                   8663
Seq. No.
                   13628 1.R1040
Contig ID
                   jC-gmro02910070d03a1
5'-most EST
                   BLASTX
Method
                   q1084415
NCBI GI
                   439
BLAST score
                   5.0e-43
E value
                   245
Match length
                   40
% identity
                   RNA-binding protein - Wood tobacco >gi_624925_dbj_BAA05170_
NCBI Description
                   (D26182) RNA-binding glycine rich protein (RGP-2)
                    [Nicotiana sylvestris]
                   8664
 Seq. No.
                   13636 1.R1040
 Contig ID
                   uC-gmronoir016d01b1
 5'-most EST
                   BLASTX
 Method
                   q2832642
 NCBI GI
                    144
 BLAST score
                    1.0e-08
 E value
                    100
 Match length
 % identity
                   (AL021710) putative protein [Arabidopsis thaliana]
 NCBI Description
                    8665
 Seq. No.
                    13638 1.R1040
 Contig ID
                    leu701150379.hl
 5'-most EST
                    BLASTX
 Method
                    q2213536
 NCBI GI
                    650
 BLAST score
                    7.0e-68
 E value
                    206
 Match length
                    66
 % identity
 NCBI Description (X98738) DNA-binding protein PD1 [Pisum sativum]
```



Contig ID 13639 1.R1040

5'-most EST LIB3170-003-Q1-K1-H7

Seq. No. 866

Contig ID 13642 1.R1040

5'-most EST LIB3170-003-Q1-K1-H8

Seq. No. 8668

Contig ID 13652 1.R1040

5'-most EST LIB3040-015-Q1-E1-H9

Seq. No. 8669

Contig ID 13658 1.R1040

5'-most EST jC-gmf102220130c11a1

Method BLASTX
NCBI GI g1929998
BLAST score 295
E value 1.0e-26
Match length 101
% identity 58

NCBI Description (U77463) NADPH-dependent HC-toxin reductase [Hordeum

vulgare]

Seq. No. 8670

Contig ID 13666_1.R1040

5'-most EST LIB3170-028-Q1-K1-G7

Seq. No. 8671

Contig ID 13666_2.R1040

5'-most EST LIB3040-013-Q1-E1-F1

Seq. No. 8672

Contig ID 13666_3.R1040 5'-most EST kl1701211343.h1

Seq. No. 8673

Contig ID 13667_1.R1040

5'-most EST LIB3040-015-Q1-E1-F9

Method BLASTX
NCBI GI g3122873
BLAST score 239
E value 5.0e-20
Match length 103
% identity 44

NCBI Description TRANSCRIPTION INITIATION PROTEIN SPT4 HOMOLOG 1 >gi_1209779

(U43923) similar to Saccharomyces cerevisiae Spt4; protein has potential N-terminal zinc-finger [Homo sapiens] >gi_1401053 (U38818) SUPT4H [Homo sapiens] >gi_1401055 (U38817) SUPT4H [Homo sapiens] >gi_1401066 (U43154) Supt4h [Mus musculus] >gi_3779194 (U96809) chromatin structural

protein homolog [Mus musculus]

>gi_4507311_ref_NP_003159.1_pSUPT4H1_ suppressor of Ty

(S.cerevisiae) 4 homolog

Seq. No. 8674



Contig ID 13674_1.R1040 5'-most EST uaw700663720.h1 Method BLASTX

NCBI GI g544150
BLAST score 106
E value 1.0e-10
Match length 157
% identity 32

NCBI Description DEK PROTEIN >gi_284375_pir__S26059 probable transforming protein (dek) - human >gi_30503_emb_CAA45536_ (X64229)

putative oncogene [Homo sapiens] >gi 4007163 emb CAA21138_

(AL031774) dek (putative oncogene) [Homo sapiens] >gi_4503249_ref_NP_003463.1_pD6S231E_DEK gene

Seq. No. 8675

Contig ID 13674 2.R1040

5'-most EST jC-gmle01810030d12d1

Seq. No. 8676

Contig ID 13676 1.R1040

5'-most EST LIB3040-015-Q1-E1-D10

Method BLASTN
NCBI GI g2924257
BLAST score 92
E value 3:0e-44
Match length 222
% identity 85

NCBI Description Tobacco chloroplast genome DNA

Seq. No. 8677

Contig ID 13681_1.R1040

5'-most EST LIB3040-015-Q1-E1-D6

Seq. No. 8678

Contig ID 13684 1.R1040

5'-most EST LIB3040-015-Q1-E1-D9

Seq. No. 8679

Contig ID 13686_1.R1040

5'-most EST LIB3040-015-Q1-E1-E10

Method BLASTX
NCBI GI g2392769
BLAST score 281
E value 7.0e-25
Match length 94
% identity 60

NCBI Description (AC002534) putative histone deacetylase [Arabidopsis

thaliana]

Seq. No. 8680

Contig ID 13687 1.R1040 5'-most EST zhf700960830.h1

Method BLASTX
NCBI GI g3341685
BLAST score 403
E value 4.0e-39
Match length 134



% identity 64
NCBI Description (AC003672) unknown protein [Arabidopsis thaliana]

Seq. No. 8681

Contig ID 13720 1.R1040

5'-most EST LIB3170-002-Q1-K1-G6

Seq. No. 8682

Contig ID 13724_1.R1040 5'-most EST gsv701047493.h1

Method BLASTX
NCBI GI g170753
BLAST score 753
E value 7.0e-80
Match length 185
% identity 74

NCBI Description (M95819) initiation factor (iso) 4F p28 subunit [Triticum

aestivuml

Seq. No. 8683

Contig ID 13730 1.R1040

5'-most EST LIB3049-002-Q1-E1-F10

Method BLASTX
NCBI GI g2262102
BLAST score 255
E value 2.0e-21
Match length 202

% identity 39

NCBI Description (AC002343) hypothetical protein [Arabidopsis thaliana]

Seq. No.

Contig ID 13734 1.R1040

5'-most EST LIB3051-109-Q1-K1-G3

8684

Method BLASTX
NCBI GI g1174779
BLAST score 1918
E value 0.0e+00
Match length 449
% identity 83

NCBI Description TRYPTOPHAN SYNTHASE BETA CHAIN 2 PRECURSOR >gi_166894

(M81620) tryptophan synthase beta-subunit [Arabidopsis thaliana] >qi 4490703 emb CAB38837.1 (AL035680) tryptophan

1/40

synthase beta-subunit (TSB2) [Arabidopsis thaliana]

Seq. No. 8685

Contig ID 13734 2.R1040 5'-most EST ssr700560895.h1

Method BLASTX
NCBI GI g1174779
BLAST score 200
E value 8.0e-16
Match length 72
% identity 60

NCBI Description TRYPTOPHAN SYNTHASE BETA CHAIN 2 PRECURSOR >gi_166894

(M81620) tryptophan synthase beta-subunit [Arabidopsis

thaliana] >gi_4490703_emb_CAB38837.1_ (AL035680) tryptophan

synthase beta-subunit (TSB2) [Arabidopsis thaliana]



```
8686
Seq. No.
Contig ID
                  13736 1.R1040
5'-most EST
                  LIB3170-017-Q1-K1-B2
Method
                  BLASTX
NCBI GI
                  g1524370
BLAST score
                  526
E value
                  1.0e-66
Match length
                  191
% identity
                  65
                  (X92491) TOM20 [Solanum tuberosum]
NCBI Description
Seq. No.
                  8687
                  13750 1.R1040
Contig ID
5'-most EST
                  jC-gmro02910060f05d1
Seq. No.
                  8688
Contig ID
                  13753 1.R1040
5'-most EST
                  LIB3040-014-Q1-E1-D9
Method
                  BLASTN
                  q3599418
NCBI GI
BLAST score
                  80
E value
                  8.0e-37
Match length
                  322
% identity
                  88
NCBI Description
                  Glycine max alternative oxidase precursor (Aox1) gene,
                  nuclear gene encoding mitochondrial protein, complete cds
                  8689
Seq. No.
Contig ID
                  13755 1.R1040
                  qsf700698435.h1
5'-most EST
```

BLASTN Method NCBI GI q1063683 BLAST score 34 E value 2.0e-09

Match length 46 93 % identity

Arabidopsis thaliana glycine-rich protein (AtGRP2b) mRNA, NCBI Description

complete cds

8690 Seq. No.

13758 1.R1040 Contig ID

5'-most EST LIB3073-005-Q1-K1-B6

BLASTX Method NCBI GI q3510256 BLAST score 352 5.0e-33 E value Match length 151 49 % identity

NCBI Description (AC005310) unknown protein [Arabidopsis thaliana]

8691 Seq. No.

13758 2.R1040 Contig ID

5'-most EST LIB3040-014-Q1-E1-E2

Method BLASTX NCBI GI g3510256 BLAST score 278



E value 2.0e-24 Match length 103 % identity

NCBI Description (AC005310) unknown protein [Arabidopsis thaliana]

Seq. No. 8692

Contig ID 13758 4.R1040 5'-most EST pcp700995143.h1

Method BLASTX NCBI GI g3510256 BLAST score 227 1.0e-18 E value Match length 112 % identity 51

(AC005310) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

8693

13773 1.R1040 Contig ID

5'-most EST LIB3170-004-Q1-K1-G7

8694 Seq. No.

13779_1.R1040 Contig ID 5'-most EST asn701141901.h1

Seq. No.

Contig ID 13780 1.R1040

5'-most EST LIB3170-001-Q1-K1-F11

8695

8696 Seq. No.

13781_1.R1040 Contig ID

5'-most EST LIB3040-013-Q1-E1-F9

8697 Seq. No.

13785 1.R1040 Contig ID

5'-most EST LIB3138-075-P1-N1-B4

Method BLASTX NCBI GI q4218124 BLAST score 176 E value 1.0e-12 Match length 44 % identity 77

NCBI Description (AL035353) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8698

13786 1.R1040 Contig ID

5'-most EST LIB3040-013-Q1-E1-H12

Method BLASTN NCBI GI g1888556 BLAST score 175 1.0e-93 E value Match length 295 % identity 90

NCBI Description Lupinus angustifolius diadenosine

5',5'''-P1,P4-tetraphosphate hydrolase mRNA, complete cds

Seq. No. 8699

Contig ID 13790 1.R1040



5'-most EST jC-gmro02910074h12a1

Method BLASTN
NCBI GI g2345147
BLAST score 288
E value 1.0e-161
Match length 447
% identity 92

NCBI Description Pisum sativum developmentally regulated GTP binding protein

(PsDRG1) mRNA, complete cds

Seq. No. 8700

Contig ID 13792 1.R1040

5'-most EST LIB3170-002-Q1-K1-F9

Seq. No. 8701

Contig ID 13801_1.R1040 5'-most EST eep700865294.h1

Method BLASTX
NCBI GI g4115377
BLAST score 1076
E value 1.0e-118
Match length 268
% identity 74

NCBI Description (AC005967) unknown protein [Arabidopsis thaliana]

Seq. No. 8702

Contig ID 13812 1.R1040

5'-most EST LIB3107-068-Q1-K1-A5

Method BLASTN
NCBI GI g343348
BLAST score 384
E value 0.0e+00
Match length 450
% identity 98

NCBI Description Glycine max mitochondrial DNA sequence, transcription

initiation motif

Seq. No. 8703

Contig ID 13813_1.R1040 5'-most EST vzy700754146.h1

Seq. No. 8704

Contig ID 13823_1.R1040

5'-most EST LIB3170-009-Q1-J1-F9

Seq. No. 8705

Contig ID 13825_1.R1040 5'-most EST fua701042577.h1

Seq. No. 8706

Contig ID 13827 1.R1040

5'-most EST LIB3056-013-Q1-N1-G10

Method BLASTX
NCBI GI g2511590
BLAST score 928
E value 1.0e-100
Match length 198



% identity

(Y13692) multicatalytic endopeptidase complex, proteasome component, beta subunit [Arabidopsis thaliana] >gi_3421111 NCBI Description

(AF043534) 20S proteasome beta subunit PBD1 [Arabidopsis

thaliana]

Seq. No.

8707

Contig ID 5'-most EST 13828 1.R1040 jex700909345.hl

Seq. No.

8708

Contig ID

13828 2.R1040

5'-most EST

jC-gmst02400032f03a1

Seq. No.

8709

Contig ID 13828 3.R1040

5'-most EST

LIB3040-013-Q1-E1-D2

Seq. No.

8710

Contig ID 5'-most EST 13829 1.R1040

Method

zhf700960810.h1 BLASTX

NCBI GI

g2673910

BLAST score

147

E value

5.0e-09

Match length

39

% identity

74

NCBI Description

(AC002561) hypothetical protein [Arabidopsis thaliana]

Seq. No.

8711

Contig ID

13830 1.R1040

5'-most EST

LIB3170-002-Q1-K1-F3

Seq. No.

8712

Contig ID

13838 1.R1040

5'-most EST Method

rry700808304.h1 BLASTX

NCBI GI

g3335366

BLAST score

270

E value

1.0e-23

Match length

91 49

% identity NCBI Description

(AC003028) unknown protein [Arabidopsis thaliana]

Seq. No.

8713

Contig ID

13847 1.R1040

5'-most EST

LIB3170-002-Q1-K1-F5

Seq. No.

8714

Contig ID

13853 1.R1040

5'-most EST

jC-gmst02400049q08a1

Seq. No.

8715

8716

Contig ID 5'-most EST 13853 2.R1040

qsv701052150.hl

Seq. No.



Contig ID 13858_1.R1040

5'-most EST LIB3040-012-Q1-E1-F5

Method BLASTX
NCBI GI g3142294
BLAST score 1676
E value 0.0e+00
Match length 374
% identity 91

NCBI Description (AC002411) Strong similarity to initiation factor eIF-2, gb_U37354 from S. pombe. ESTs gb_T41979, gb_N37284 and

gb N37529 come from this gene. [Arabidopsis thaliana]

Seq. No. 8717

Contig ID 13858_2.R1040 5'-most EST k11701205184.h1

Method BLASTX
NCBI GI g2832707
BLAST score 227
E value 8.0e-19
Match length 56
% identity 84

NCBI Description (AL021713) translation initiation factor eIF-2 gamma

chain-like protein [Arabidopsis thaliana]

Seq. No. 8718

Contig ID 13863 1.R1040

5'-most EST uC-gmflminsoy003c10b1

Method BLASTX
NCBI GI g4008441
BLAST score 458
E value 2.0e-45
Match length 143
% identity 62

NCBI Description (AL034488) predicted using Genefinder; cDNA EST yk433c6.3

comes from this gene; cDNA EST EMBL:D72601 comes from this gene; cDNA EST EMBL:D75524 comes from this gene; cDNA EST yk433c6.5 comes from this gene [Caenorhabditis elegans]

Seq. No. 8719

Contig ID 13863 2.R1040 5'-most EST ckk700605883.h2

Method BLASTX
NCBI GI g4008441
BLAST score 212
E value 5.0e-28
Match length 100
% identity 65

NCBI Description (AL034488) predicted using Genefinder; cDNA EST yk433c6.3

comes from this gene; cDNA EST EMBL:D72601 comes from this gene; cDNA EST EMBL:D75524 comes from this gene; cDNA EST yk433c6.5 comes from this gene [Caenorhabditis elegans]

Seq. No. 8720

Contig ID 13863 3.R1040

5'-most EST LIB3051-002-Q1-E1-F4

Method BLASTX NCBI GI 94008441



BLAST score 340 E value 6.0e-32 Match length 105 % identity 62

(AL034488) predicted using Genefinder; cDNA EST yk433c6.3 comes from this gene; cDNA EST EMBL:D72601 comes from this gene; cDNA EST EMBL:D75524 comes from this gene; cDNA EST yk433c6.5 comes from this gene [Caenorhabditis elegans]

Seq. No. 8721

NCBI Description

Contig ID 13868 1.R1040 5'-most EST zzp700835456.h1

Seq. No. 8722

Contig ID 13872 1.R1040 5'-most EST g4313489

Seq. No. 8723

Contig ID 13872_2.R1040 5'-most EST kl1701207410.hl

Seq. No. 8724

Contig ID 13872_3.R1040

5'-most EST uC-gmrominsoy194h08b1

Seq. No. 8725

Contig ID 13873 1.R1040

5'-most EST LIB3040-012-Q1-E1-C1

Seq. No. 8726

Contig ID 13877_1.R1040

5'-most EST LIB3040-012-Q1-E1-C5

Method BLASTN
NCBI GI g1150683
BLAST score 105
E value 7.0e-52
Match length 220
% identity 92

NCBI Description V.radiata atpB, rbcL and trnK genes

Seq. No. 8727

Contig ID 13884_1.R1040 5'-most EST bth700843663.h1

Method BLASTX
NCBI GI 94539351
BLAST score 181
E value 2.0e-13
Match length 101
% identity 46

NCBI Description (AL035539) putative protein [Arabidopsis thaliana]

Seq. No. 8728

Contig ID 13884 2.R1040

5'-most EST jC-qmro02910004e02a1

Method BLASTX
NCBI GI g4539351
BLAST score 278



E value 1.0e-24 Match length 107 % identity 61

NCBI Description (AL035539) putative protein [Arabidopsis thaliana]

Seq. No. 8729

Contig ID 13889 1.R1040

5'-most EST LIB3170-004-Q1-K1-F1

Seq. No. 8730

Contig ID 13891 1.R1040

5'-most EST LIB3040-011-Q1-E1-H5

Method BLASTX
NCBI GI g2245124
BLAST score 734
E value 1.0e-77
Match length 241
% identity 65

NCBI Description (Z97343) protein kinase [Arabidopsis thaliana]

Seq. No. 8731

Contig ID 13891 3.R1040

5'-most EST uC-gmflminsoy008c09b1

Method BLASTX
NCBI GI g2245124
BLAST score 167
E value 1.0e-11
Match length 74
% identity 53

NCBI Description (Z97343) protein kinase [Arabidopsis thaliana]

Seq. No. 8732

Contig ID 13900_1.R1040 5'-most EST fde700876566.h1

Method BLASTX
NCBI GI g4522012
BLAST score 706
E value 2.0e-74
Match length 234
% identity 58

NCBI Description (AC007069) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8733

Contig ID 13902 1.R1040

5'-most EST LIB3170-002-Q1-K1-E3

Seq. No. 8734

Contig ID 13903 1.R1040

5'-most EST LIB3170-001-Q1-K1-E4

Seq. No. 8735

Contig ID 13906_1.R1040 5'-most EST asn701134208.h1

Method BLASTX
NCBI GI g349379
BLAST score 342
E value 3.0e-39



Match length 134 % identity 65

NCBI Description (L22847) HAHB-1 [Helianthus annuus]

Seq. No.

8736

Contig ID 13907 1.R1040

5'-most EST LIB3170-002-Q1-K1-E5

Seq. No.

8737

Contig ID

13922 1.R1040

5'-most EST

LIB3040-011-Q1-E1-B3

Seq. No.

8738

Contig ID

13926_1.R1040

5'-most EST

LIB3040-011-Q1-E1-B8

Seq. No.

8739

Contig ID

13928 1.R1040

5'-most EST

LIB3170-002-Q1-K1-E2

Seq. No.

8740

Contig ID 5'-most EST 13930_1.R1040 leu701152715.h1

Seq. No.

8741

Contig ID

13931 1.R1040

5'-most EST

LIB3074-019-Q1-E1-H1

Method NCBI GI BLASTX q2407233

BLAST score

92407233 280

E value

2.0e-24

Match length

101

% identity

48

NCBI Description

(AF017750) similar to Haemophilus influenzae product

encoded by Genbank Accession Number U32796 [Haemophilus

ducreyi]

Seq. No.

8742

Contig ID

13931 2.R1040

5'-most EST

jC-gmle01810043e07a1

Seq. No.

8743

Contig ID

13931 4.R1040

5'-most EST

hrw701061635.h1

Seq. No.

8744

Contig ID

13935_1.R1040

5'-most EST

fC-gmse700671588a3

Method NCBI GI BLASTX g2583129

BLAST score

1833

E value

1000

Match length

0.0e+00

% identity

394 83

NCBI Description

(AC002387) putative methionine aminopeptidase [Arabidopsis

thaliana]



Contig ID 13936 1.R1040

5'-most EST LIB3049-034-Q1-E1-F3

Method BLASTX
NCBI GI g4539457
BLAST score 327
E value 4.0e-60
Match length 278
% identity 53

NCBI Description (AL049500) heat shock transcription factor-like protein

[Arabidopsis thaliana]

Seq. No. 8746

Contig ID 13936 2.R1040

5'-most EST LIB3093-034-Q1-K1-D8

Method BLASTX
NCBI GI g4539457
BLAST score 600
E value 7.0e-62
Match length 237
% identity 57

NCBI Description (AL049500) heat shock transcription factor-like protein

[Arabidopsis thaliana]

Seq. No. 8747

Contig ID 13937 1.R1040

5'-most EST LIB3170-004-Q1-K1-E7

Method BLASTX
NCBI GI g3873408
BLAST score 200
E value 3.0e-15
Match length 90
% identity 41

NCBI Description (L76926) putative zinc finger protein [Arabidopsis

thaliana]

Seq. No. 8748

Contig ID 13954_1.R1040

5'-most EST jC-gmst02400074e06a1

Seq. No. 8749

Contig ID 13973_1.R1040

5'-most EST LIB3170-001-Q1-K1-D8

Method BLASTX
NCBI GI g2760323
BLAST score 1146
E value 1.0e-126
Match length 284
% identity 73

NCBI Description (AC002130) F1N21.8 [Arabidopsis thaliana]

Seq. No. 8750

Contig ID 13974_1.R1040

5'-most EST LIB3170-002-Q1-J1-D8

Seq. No. 8751

Contig ID 13989 1.R1040



5'-most EST LIB3170-003-Q1-J1-E2

Method BLASTN
NCBI GI g3599418
BLAST score 104
E value 7.0e-52
Match length 135
% identity 94

NCBI Description Glycine max alternative oxidase precursor (Aox1) gene,

nuclear gene encoding mitochondrial protein, complete cds

Seq. No. Contig ID

8752 14001 1.R1040

5'-most EST LI

LIB3094-062-Q1-K1-C7

Method BLASTN NCBI GI g2062691 BLAST score 34

E value 3.0e-09 Match length 34

% identity 62

NCBI Description Human sodium phosphate transporter (NPT4) mRNA, complete

cds

Seq. No.

8753

Contig ID 14001 2.R1040

5'-most EST

LIB3040-010-Q1-E1-A10

Seq. No.

8754

Contig ID 14004 1.R1040

5'-most EST

LIB3170-001-Q1-J1-D4

Seq. No.

8755

Contig ID 14013_1.R1040

5'-most EST

LIB3049-039-Q1-E1-D11

Seq. No.

8756

Contig ID 14047_1.R1040

5'-most EST LIB3050-024-Q1-K1-C5

Method BLASTX
NCBI GI g3236242
BLAST score 428
E value 6.0e-42
Match length 103
% identity 84

NCBI Description (AC004684) putative ribosomal protein L36 [Arabidopsis

thaliana]

Seq. No.

8757

Contig ID

14047_2.R1040

5'-most EST

LIB3093-027-Q1-K1-H11

Method BLASTX
NCBI GI g3236242
BLAST score 407
E value 1.0e-39
Match length 99
% identity 83

NCBI Description (AC004684) putative ribosomal protein L36 [Arabidopsis

thaliana]



Contig ID 14047 3.R1040

5'-most EST LIB3049-049-Q1-E1-E6

Method BLASTX
NCBI GI g1710546
BLAST score 437
E value 4.0e-43
Match length 106
% identity 79

NCBI Description 60S RIBOSOMAL PROTEIN L36 >gi_1276967 (U47095) putative

ribosomal protein [Daucus carota]

Seq. No. 8759

Contig ID 14047_4.R1040 5'-most EST vwf700675179.h1

Method BLASTX
NCBI GI g1710546
BLAST score 346
E value 6.0e-33
Match length 75
% identity 89

NCBI Description 60S RIBOSOMAL PROTEIN L36 >gi 1276967 (U47095) putative

ribosomal protein [Daucus carota]

Seq. No. 8760

Contig ID 14049 1.R1040

5'-most EST LIB3040-009-Q1-E1-B11

Seq. No. 8761

Contig ID 14057 1.R1040

5'-most EST LIB3040-009-Q1-E1-C5

Seq. No. 8762

Contig ID 14058_1.R1040 5'-most EST eep700870490.h1

Method BLASTX
NCBI GI g1532168
BLAST score 251
E value 2.0e-21
Match length 66
% identity 73

NCBI Description (U63815) localized according to blastn similarity to EST

sequences; therefore, the coding span corresponds only to an area of similarity since the initation codon and stop codon could not be precisely determined [Arabidopsis

thaliana]

Seq. No. 8763

Contig ID 14066 1.R1040

5'-most EST LIB3170-006-Q1-K1-H7

Method BLASTX
NCBI GI g2829910
BLAST score 440
E value 2.0e-43
Match length 234
% identity 45



NCBI Description (AC002291) Unknown protein, contains regulator of chromosome condensation motifs [Arabidopsis thaliana]

Seq. No. 8764

Contig ID 14068 1.R1040

5'-most EST LIB3167-023-P4-K4-E8

Seq. No. 8765

Contig ID 14076 1.R1040 5'-most EST ncj700988351.h1

Method BLASTX
NCBI GI g3342798
BLAST score 453
E value 7.0e-45
Match length 132
% identity 64

NCBI Description (AF061240) glutamine cyclotransferase precursor [Carica

papaya]

Seq. No. 8766

Contig ID 14076_2.R1040
5'-most EST g5606901
Method BLASTX
NCBI GI g3342798
BLAST score 251
E value 6.0e-27
Match length 166

% identity 53

NCBI Description (AF061240) glutamine cyclotransferase precursor [Carica

papaya]

Seq. No. 8767

Contig ID 14082 1.R1040

5'-most EST LIB3049-007-Q1-E1-E9

Seq. No. 8768

Contig ID 14082 2.R1040

5'-most EST LIB3040-008-Q1-E1-F5

Seq. No. 8769

Contig ID 14086 1.R1040

5'-most EST LIB3049-009-Q1-E1-F4

Method BLASTX
NCBI GI g3881507
BLAST score 461
E value 2.0e-45
Match length 147
% identity 57

NCBI Description (Z47357) cDNA EST yk375c3.5 comes from this gene; cDNA EST

yk375c3.3 comes from this gene [Caenorhabditis elegans]

Seq. No. 8770

Contig ID 14086 2.R1040 5'-most EST gsv701046761.h1

Method BLASTX NCBI GI g3881507 BLAST score 196



E value 5.0e-15
Match length 53
% identity 60
NCBI Description (Z47357

BI Description (Z47357) cDNA EST yk375c3.5 comes from this gene; cDNA EST yk375c3.3 comes from this gene [Caenorhabditis elegans]

Seq. No. 8771

Contig ID 14092 1.R1040

5'-most EST LIB3170-039-Q1-K2-F2

Method BLASTX
NCBI GI g2924520
BLAST score 754
E value 6.0e-80
Match length 273
% identity 82

NCBI Description (AL022023) plasma membrane intrinsic protein (SIMIP)

[Arabidopsis thaliana]

Seq. No. 8772

Contig ID 14096 1.R1040

5'-most EST LIB3106-067-P1-K1-C6

Method BLASTX
NCBI GI g2618688
BLAST score 430
E value 2.0e-42
Match length 99
% identity 77

NCBI Description (AC002510) putative esterase D [Arabidopsis thaliana]

Seq. No. 8773

Contig ID 14104 1.R1040 5'-most EST gsv701049590.h1

Method BLASTX
NCBI GI g2194132
BLAST score 485
E value 7.0e-49
Match length 156
% identity 59

NCBI Description (AC002062) No definition line found [Arabidopsis thaliana]

Seq. No. 8774

Contig ID 14105_1.R1040 5'-most EST pxt700945486.h1

Method BLASTX
NCBI GI 94314389
BLAST score 326
E value 3.0e-30
Match length 91
% identity 68

NCBI Description (AC006232) putative transcription factor [Arabidopsis

thalianal

Seq. No. 8775

Contig ID 14114_1.R1040

5'-most EST LIB3170-002-Q1-K1-C4

Seq. No. 8776



Contig ID 14116_1.R1040 5'-most EST LIB3051-067-Q1-K1-B5

Method BLASTN
NCBI GI g2351065
BLAST score 43

E value 1.0e-14
Match length 130
% identity 74

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MHF15, complete sequence [Arabidopsis thaliana]

Seq. No. 8777

Contig ID 14120_1.R1040

5'-most EST LIB3107-017-Q1-K1-C12

Seq. No. 8778

Contig ID 14120 2.R1040 5'-most EST g4307297

Method BLASTX
NCBI GI g1938549
BLAST score 203
E value 1.0e-15
Match length 122

Match length 122 % identity 42

NCBI Description (U97016) similar to drosophila Rlc1 gene product

(NID:g563361) and S. cerevisiae mitochondrial 60S ribosomal protein L4 (YML4) (NID:g459259) [Caenorhabditis elegans]

Seq. No. 8779

Contig ID 14122 1.R1040

5'-most EST jC-gmro02910011b12a1

Method BLASTX
NCBI GI g2583120
BLAST score 262
E value 2.0e-40
Match length 143
% identity 47

NCBI Description (AC002387) putative receptor-like protein kinase

[Arabidopsis thaliana]

Seq. No. 8780

Contig ID 14128_1.R1040

5'-most EST LIB3049-032-Q1-E1-D9

Seq. No. 8781

Contig ID 14137_1.R1040

5'-most EST LIB3040-008-Q1-E1-A2

Seq. No. 8782

Contig ID 14137 2.R1040

5'-most EST LIB3106-105-Q1-K1-E2

Seq. No. 8783

Contig ID 14146_1.R1040

5'-most EST LIB3049-025-Q1-E1-B1

Method BLASTX NCBI GI g3184100



BLAST score 270 E value 2.0e-23 Match length 72 % identity 67

NCBI Description (AL023777) rna binding protein [Schizosaccharomyces pombe]

Seq. No. 8784

Contig ID 14146 2.R1040 5'-most EST asn701142530.h1

Seq. No. 8785

Contig ID 14164_1.R1040

5'-most EST LIB3040-007-Q1-E1-A8

Method BLASTX
NCBI GI g3269289
BLAST score 371
E value 2.0e-35
Match length 112
% identity 66

NCBI Description (AL030978) putative protein [Arabidopsis thaliana]

Seq. No. 8786

Contig ID 14164 2.R1040 5'-most EST zsg701123057.h1

Method BLASTX
NCBI GI g3269289
BLAST score 369
E value 3.0e-35
Match length 137
% identity 69

NCBI Description (AL030978) putative protein [Arabidopsis thaliana]

Seq. No. 8787

Contig ID 14174 1.R1040

5'-most EST LIB3073-026-Q1-K1-G11

Seq. No. 8788

Contig ID 14174_2.R1040

5'-most EST LIB3170-004-Q1-K1-C1

Seq. No. 8789

Contig ID 14176_1.R1040

5'-most EST LIB3170-003-Q1-K1-C1

Seq. No. 8790

Contig ID 14204 1.R1040 5'-most EST zpv700760194.h1

Method BLASTX
NCBI GI g3367536
BLAST score 521
E value 7.0e-53
Match length 117
% identity 88

NCBI Description (AC004392) Contains similarity to symbiosis-related like

protein F1N20.80 gi_2961343 from A. thaliana BAC gb AL022140. EST gb T04695 comes from this gene.

[Arabidopsis thaliana]



Contig ID 14204 2.R1040

5'-most EST LIB3109-050-Q1-K1-C4

Method BLASTX
NCBI GI g3367536
BLAST score 520
E value 6.0e-53
Match length 117
% identity 88

NCBI Description (AC004392) Contains similarity to symbiosis-related like

protein F1N20.80 gi_2961343 from A. thaliana BAC gb_AL022140. EST gb_T04695 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 8792

Contig ID 14204 3.R1040 5'-most EST epx701109088.h1

Method BLASTX
NCBI GI g3367536
BLAST score 316
E value 4.0e-29
Match length 85
% identity 79

NCBI Description (AC004392) Contains similarity to symbiosis-related like

protein F1N20.80 gi_2961343 from A. thaliana BAC gb AL022140. EST gb_T04695 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 8793

Contig ID 14205_1.R1040

5'-most EST LIB3170-004-Q1-K1-B8

Method BLASTX
NCBI GI g2708743
BLAST score 222
E value 4.0e-18
Match length 142
% identity 32

NCBI Description (AC003952) putative Tal-1-like reverse transcriptase

[Arabidopsis thaliana]

Seq. No. 8794

Contig ID 14212_1.R1040

5'-most EST LIB3040-006-Q1-E1-F7

Seq. No. 8795

Contig ID 14215_1.R1040 5'-most EST zzp700831774.h1

Method BLASTX
NCBI GI g4510386
BLAST score 550
E value 4.0e-56
Match length 161
% identity 64

NCBI Description (AC007017) unknown protein [Arabidopsis thaliana]

Seq. No. 8796



```
14215 2.R1040
Contig ID
                  LIB3170-002-Q1-K1-A10
5'-most EST
                  BLASTX
Method
                  g4510387
NCBI GI
                  396
BLAST score
                  2.0e-38
E value
                  122
Match length
% identity
                   62
NCBI Description (AC007017) unknown protein [Arabidopsis thaliana]
                   8797
Seq. No.
                   14218 1.R1040
Contig ID
                   LIB3040-006-Q1-E1-C1
5'-most EST
                   8798
Seq. No.
                   14224 1.R1040
Contig ID
                   LIB3139-052-P1-N1-A5
5'-most EST
                   8799
Seq. No.
                   14227 1.R1040
Contig ID
                   LIB3040-006-Q1-E1-D12
5'-most EST
                   8800
Seq. No.
                   14230 1.R1040
Contig ID
                   LIB3040-005-Q1-E1-H5
5'-most EST
                   BLASTX
Method
                   g3377813
NCBI GI
                   328
BLAST score
                   3.0e-30
E value
                   150
Match length
                   49
% identity
                   (AF076275) No definition line found [Arabidopsis thaliana]
NCBI Description
                   8801
Seq. No.
                   14231 1.R1040
Contig ID
                   zhf700953828.hl
 5'-most EST
                   BLASTX
Method
                   g3063442
NCBI GI
BLAST score
                   181
                   6.0e-13
E value
Match length
                   84
 % identity
                   (AC003981) F22013.6 [Arabidopsis thaliana]
 NCBI Description
                   8802
 Seq. No.
                   14233 1.R1040
 Contig ID
                   LIB3040-005-Q1-E1-H9
 5'-most EST
 Method
                   BLASTX
                   q3377813
 NCBI GI
                    426
 BLAST score
```

1.0e-41 E value 202 Match length % identity

(AF076275) No definition line found [Arabidopsis thaliana] NCBI Description

Seq. No.

14240 1.R1040 Contig ID



pmv700889748.h1 5'-most EST

BLASTX Method g4455754 NCBI GI BLAST score 164 6.0e-11 E value Match length 160 % identity 28

(AL035478) hypothetical protein SC2G5.30 [Streptomyces NCBI Description

coelicolor]

8804 Seq. No.

14242 1.R1040 Contig ID

5'-most EST jC-gmro02910025g01d1

Method BLASTX NCBI GI g4220485 BLAST score 455 E value 4.0e-45 Match length 139 % identity

(AC006069) putative beta-1,3-glucanase [Arabidopsis NCBI Description

thaliana]

Seq. No. 8805

14245 1.R1040 Contig ID

LIB3170-002-Q1-K1-A11 5'-most EST

8806 Seq. No.

Contig ID 14251 1.R1040

jC-gmst02400071e04a1 5'-most EST

Method BLASTX q4006827 NCBI GI BLAST score 686 E value 2.0e-72 182 Match length % identity

(AC005970) subtilisin-like protease [Arabidopsis thaliana] NCBI Description

Seq. No. 8807

Contig ID 14252 1.R1040

5'-most EST LIB3051-037-Q1-K1-C6

Method BLASTN NCBI GI g3821780 BLAST score 37 5.0e-11 E value Match length 37 100 % identity

NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No.

14252 2.R1040 Contig ID

5'-most EST LIB3040-005-Q1-E1-F9

8808

8809 Seq. No.

14282 1.R1040 Contig ID

5'-most EST LIB3040-005-Q1-E1-A12

Seq. No. 8810

1583

Match length

184



```
Contig ID
                  14283 1.R1040
                                                                       ÷,
                  LIB3170-003-Q1-K1-A1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3645899
BLAST score
                  173
E value
                  1.0e-12
                  60
Match length
                  58
% identity
NCBI Description (U68408) 5' end not determined experimentally [Zea mays]
                  8811
Seq. No.
                  14299 1.R1040
Contig ID
5'-most EST
                  txt700733446.h1
Seq. No.
                  8812
                  14300 1.R1040
Contig ID
5'-most EST
                  LIB3170-018-Q1-K1-G10
                  8813
Seq. No.
Contig ID
                  14302 1.R1040
5'-most EST
                  LIB3170-086-Q1-J1-B7
Method
                  BLASTX
NCBI GI
                  q2739387
BLAST score
                  225
E value
                  4.0e-18
Match length
                  116
% identity
                  43
NCBI Description (AC002505) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                  8814
                  14306 1.R1040
Contig ID
5'-most EST
                  jC-gmf102220094a10a1
Method
                  BLASTX
NCBI GI
                  q1173121
BLAST score
                  217
E value
                  4.0e-17
Match length
                  116
% identity
                  44
NCBI Description ATP-DEPENDENT RNA HELICASE ROK1 >gi 2131186 pir S59649
                  ATP-dependent RNA helicase ROK1 - yeast (Saccharomyces
                  cerevisiae) >gi 607182 emb CAA84384 (Z34901) ATP-dependent
                  RNA helicase [Saccharomyces cerevisiae]
                  >gi 971382 emb CAA59758 (X85757) putative ATP dependent
                  RNA helicase [Saccharomyces cerevisiae]
                  >gi 1322774 emb CAA96883 (Z72693) ORF YGL171w
                  [Saccharomyces cerevisiae]
                  8815
Seq. No.
                  14313 1.R1040
Contig ID
5'-most EST
                  LIB3040-004-Q1-E1-G11
Method
                  BLASTN
                  g2398828
NCBI GI
BLAST score
                  76
E value
                  2.0e-34
```

% identity 85
NCBI Description Solanum tuberosum mRNA for mitochondrial uncoupling protein



```
8816
Seq. No.
Contig ID
                  14322 1.R1040
5'-most EST
                  vzy700754492.h1
                  BLASTX
Method
NCBI GI
                  q3881978
BLAST score
                  1322
                  1.0e-146
E value
                  324
Match length
                  81
% identity
                 (Y11348) annexin-like protein [Medicago sativa]
NCBI Description
                  8817
Seq. No.
                  14322 2.R1040
Contig ID
5'-most EST
                  LIB3051-116-Q1-K1-F3
Method
                  BLASTX
NCBI GI
                  g3881978
BLAST score
                  432
E value
                  1.0e-42
Match length
                  110
% identity
                  33
                 (Y11348) annexin-like protein [Medicago sativa]
NCBI Description
Seq. No.
                  8818
Contig ID
                  14323 1.R1040
5'-most EST
                  LIB3107-044-Q1-K1-D3
Seq. No.
                  8819
                  14326 1.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy070e10b1
Method
                  BLASTX
NCBI GI
                  g2492668
BLAST score
                  275
E value
                  4.0e-24
Match length
                  176
% identity
                  39
NCBI Description ACTIN >gi 1419540_emb_CAA67388_ (X98885) beta-actin [Fucus
                  vesiculosus]
                  8820
Seq. No.
Contig ID
                  14326 2.R1040
5'-most EST
                  LIB3107-009-Q1-K1-C5
Method
                  BLASTX
NCBI GI
                  q728795
BLAST score
                  187
E value
                  1.0e-13
Match length
                  90
% identity
                  43
                  CENTRACTIN (ACTIN-LIKE PROTEIN) >gi 626122 pir S48973
NCBI Description
                  centractin ACT3 - yeast (Saccharomyces cerevisiae)
                  >gi_500679 (U10398) Arplp: Actin-related protein, of the
```

dynactin complex [Saccharomyces cerevisiae]

>gi_557670_emb_CAA56206_ (X79811) centractin [Saccharomyces

cerevisiae]

Seq. No. 8821

Contig ID 14326_3.R1040



```
5'-most EST
                   LIB3107-034-Q1-K1-H6
                   8822
Seq. No.
Contig ID
                   14326 5.R1040
5'-most EST
                   sat701010945.hl
                   8823
Seq. No.
                   14342_1.R1040
Contig ID
5'-most EST
                   jC-gmst02400042h11a1
Method
                   BLASTX
                   g2388913
NCBI GI
BLAST score
                   141
E value
                   1.0e-08
Match length
                   104
% identity
                   42
                  (Z98974) hypothetical protein [Schizosaccharomyces pombe]
NCBI Description
Seq. No.
                   8824
Contig ID
                   14344 1.R1040
5'-most EST
                   LIB3170-020-Q1-J1-G12
Seq. No.
                   8825
                   14351 1.R1040
Contig ID
5'-most EST
                   uC-gmropic083h01b1
Method
                   BLASTX
NCBI GI
                   g4544453
BLAST score
                   376
E value
                   4.0e-36
Match length
                   105
% identity
                   73
NCBI Description (AC006592) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   8826
Contig ID
                   14351 2.R1040
5'-most EST
                   LIB3170-019-Q1-K1-G10
Method
                   BLASTX
                   g4544453
NCBI GI
BLAST score
                   280
                   1.0e-31
E value
                   89
Match length
% identity
                   81
NCBI Description
                  (AC006592) hypothetical protein [Arabidopsis thaliana]
                   8827
Seq. No.
Contig ID
                   14357 1.R1040
5'-most EST
                   k1170\overline{1}209107.h1
Method
                   BLASTX
NCBI GI
                   q4521249
```

Method BLASTX
NCBI GI g4521249
BLAST score 468
E value 2.0e-46
Match length 268
% identity 46

NCBI Description (AB013912) DNA helicase [Mus musculus]

Seq. No. 8828

Contig ID 14358_1.R1040 5'-most EST LIB3040-004-Q1-E1-A1



Seq. No. 8829

Contig ID 14359 1.R1040

5'-most EST LIB3040-004-Q1-E1-A11

Method BLASTX
NCBI GI g3980034
BLAST score 286
E value 2.0e-25
Match length 81
% identity 65

NCBI Description (AL033514) predicted using Genefinder; cDNA EST EMBL:D71127

comes from this gene; cDNA EST EMBL:D73731 comes from this

gene [Caenorhabditis elegans]

Seq. No. 8830

Contig ID 14362_1.R1040 5'-most EST zhf700959092.h1

Seq. No.

Contig ID 14377 1.R1040

5'-most EST uC-gmrominsoy203b10b1

8831

Method BLASTX
NCBI GI g4455224
BLAST score 357
E value 2.0e-37
Match length 188
% identity 46

NCBI Description (AL035440) putative protein [Arabidopsis thaliana]

Seq. No. 8832

Contig ID 14382 1.R1040

5'-most EST LIB3040-003-Q1-E1-G4

Seq. No. 8833

Contig ID 14386_1.R1040 5'-most EST ncj700985656.h1

Method BLASTX
NCBI GI g2257503
BLAST score 437
E value 8.0e-43
Match length 272
% identity 38

NCBI Description (AB004535) BEM46 PROTEIN [Schizosaccharomyces pombe]

Seq. No. 8834

Contig ID 14386 2.R1040 5'-most EST kl1701213588.h1

Seq. No. 8835

Contig ID 14387 1.R1040

5'-most EST LIB3040-003-Q1-E1-D1

Seq. No. 8836

Contig ID 14398 1.R1040

5'-most EST LIB3170-079-Q1-J1-G1

Seq. No. 8837



Contig ID 14398_2.R1040

5'-most EST LIB3170-020-Q1-J1-G3

Seq. No. 8838

Contig ID 14398_3.R1040 5'-most EST jex700903540.h1

Seq. No. 8839

Contig ID 14409_1.R1040 5'-most EST eep700867225.h1

Method BLASTX
NCBI GI g2649424
BLAST score 143
E value 7.0e-09
Match length 100
% identity 31

NCBI Description (AE001023) A. fulgidus predicted coding region AF1178

[Archaeoglobus fulgidus]

Seq. No. 8840

Contig ID 14410 1.R1040

5'-most EST jC-gmle01810041c03a1

Seq. No. 8841

Contig ID 14414 1.R1040

5'-most EST LIB3107-005-Q1-K1-G11

Seq. No. 8842

Contig ID 14419_1.R1040 5'-most EST fde700872835.h1

Method BLASTX
NCBI GI g585322
BLAST score 231
E value 5.0e-19
Match length 84
% identity 51

NCBI Description INORGANIC PYROPHOSPHATASE (PYROPHOSPHATE PHOSPHO-HYDROLASE)

(PPASE) >gi 539751 pir A45153 inorganic pyrophosphatase

(EC 3.6.1.1) - bovine

Seq. No. 8843

Contig ID 14419_3.R1040 5'-most EST ssr700556335.h1

Seq. No. 8844

Contig ID 14425 1.R1040

5'-most EST LIB3040-002-Q1-E1-H3

Seq. No. 8845

Contig ID 14427_1.R1040

5'-most EST LIB3040-002-Q1-E1-H7

Seq. No. 8846

Contig ID 14432 1.R1040

5'-most EST LIB3040-003-Q1-E1-A2

Method BLASTX NCBI GI g4325354



BLAST score 244 E value 7.0e-25 Match length 96 % identity 62

NCBI Description (AF128395) contains similarity to retrovirus-related polyproteins and to CCHC zinc finger protein (Pfam:

PF00098, Score=16.3, E=0.051, E= 1) [Arabidopsis thaliana]

Seq. No. 8847

Contig ID 14434 1.R1040

5'-most EST LIB3040-002-Q1-E1-E7

Seq. No. 8848

Contig ID 14437 1.R1040

5'-most EST LIB3170-020-Q1-J1-F8

Seq. No. 8849

Contig ID 14447 1.R1040

5'-most EST LIB3040-002-Q1-E1-G11

Seq. No. 8850

Contig ID 14450 1.R1040

5'-most EST LIB3040-002-Q1-E1-C6

Method BLASTX
NCBI GI g2947062
BLAST score 248
E value 7.0e-21
Match length 168
% identity 42

NCBI Description (AC002521) unknown protein [Arabidopsis thaliana]

Seq. No. 8851

Contig ID 14457 1.R1040

5'-most EST LIB3170-079-Q1-K1-D10

Method BLASTX
NCBI GI g2134723
BLAST score 275
E value 1.0e-23
Match length 191
% identity 38

NCBI Description AlkB homolog protein ABH - human >gi 1237210 emb CAA63047

(X91992) alkB homolog protein [Homo sapiens]

>gi_1588535_prf__2208455A ABH gene [Homo sapiens]

Seq. No. 8852

Contig ID 14458 1.R1040

5'-most EST LIB3040-002-Q1-E1-D8

Method BLASTX
NCBI GI 9466044
BLAST score 153
E value 3.0e-15
Match length 146
% identity 35

NCBI Description HYPOTHETICAL ZINC FINGER PROTEIN ZK686.4 IN CHROMOSOME III

>gi_630780_pir__S44909 ZK686.4 protein - Caenorhabditis
elegans >gi 304346 (L17337) coded for by C. elegans cDNAs

GenBank: M88869 and T01933; putative [Caenorhabditis



elegans]

Seq. No. 8853

Contig ID 14463_1.R1040 5'-most EST uC-gmronoir013f02b1

Method BLASTX
NCBI GI g3024652
BLAST score 195
E value 1.0e-14
Match length 155
% identity 36

NCBI Description SORCIN >gi 1655733 (U39069) sorcin [Schistosoma japonicum]

Seq. No. 885

Contig ID 14463_3.R1040 5'-most EST leu701144523.h1

Seq. No. 8855

Contig ID 14467 1.R1040

5'-most EST LIB3040-002-Q1-E1-B10

Seq. No. 8856

Contig ID 14469 1.R1040

5'-most EST jC-gmro02910034b12a1

Method BLASTN
NCBI GI g3819163
BLAST score 586
E value 0.0e+00
Match length 638
% identity 98

NCBI Description Glycine max cctd gene

Seq. No. 8857

Contig ID 14476 1.R1040

5'-most EST LIB3040-002-Q1-E1-B8

Method BLASTX
NCBI GI g3024666
BLAST score 196
E value 4.0e-15
Match length 90
% identity 46

NCBI Description STRICTOSIDINE SYNTHASE 1/2 PRECURSOR >gi_1754983 (U43713) strictosidine synthase [Arabidopsis thaliana] >gi_1754985

(U43945) strictosidine synthase [Arabidopsis thaliana]

Seq. No. 8858

Contig ID 14481 1.R1040 5'-most EST rlr700901879.h1

Seq. No. 8859

Contig ID 14481_2.R1040 5'-most EST sat701002952.h1

Seq. No. 8860

Contig ID 14483_1.R1040

5'-most EST jC-gmf102220077f06a1

Method BLASTX



```
a2804280
NCBI GI
                  497
BLAST score
                  3.0e-50
E value
                  133
Match length
% identity
                   (AB003687) 6-4 photolyase [Arabidopsis thaliana]
NCBI Description
                  >gi_3929918_dbj_BAA34711_ (AB017331) 6-4 photolyase
                   [Arabidopsis thaliana]
                  8861
Seq. No.
                  14492 1.R1040
Contig ID
                  LIB3138-002-Q1-N1-E2
5'-most EST
                  BLASTN
Method
                  g2924257
NCBI GI
                  191
BLAST score
                   1.0e-102
E value
                   739
Match length
                   86
% identity
                  Tobacco chloroplast genome DNA
NCBI Description
                   8862
Seq. No.
                   14492 2.R1040
Contig ID
                   LIB3109-001-Q1-K1-A6
5'-most EST
                   BLASTN
Method
                   g2924257
NCBI GI
                   82
BLAST score
E value
                   7.0e-38
                   375
Match length
                   85
% identity
                   Tobacco chloroplast genome DNA
NCBI Description
                   8863
Seq. No.
                   14497 1.R1040
Contig ID
                   LIB3040-002-Q1-E1-A3
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3327269
BLAST score
                   172
                   3.0e-12
E value
                   99
Match length
                   49
% identity
                   (AB015999) PKn1 [Ipomoea nil]
NCBI Description
                   8864
Seq. No.
                   14509 1.R1040
Contig ID
                   LIB3170-018-Q1-J1-E4
5'-most EST
Seq. No.
                   8865
                   14512 1.R1040
Contig ID
                   zsq701119070.hl
 5'-most EST
                   BLASTX
Method
                   q4567312
 NCBI GI
                   325
BLAST score
```

NCBI Description (AC005956) hypothetical protein [Arabidopsis thaliana]

1.0e-29

221

44

E value

Match length

% identity



Seq. No. 8866

Contig ID 14512_2.R1040

5'-most EST LIB3051-090-Q1-K1-D3

Seq. No. 8867

Contig ID 14516_1.R1040

5'-most EST LIB3087-010-Q1-K1-C1 Method BLASTX

Method BLASTX
NCBI GI g2181184
BLAST score 672
E value 1.0e-70
Match length 180
% identity 72

NCBI Description (Y13577) JR3 protein [Arabidopsis thaliana]

Seq. No. 8868

Contig ID 14524_1.R1040

5'-most EST LIB3040-001-Q1-E1-C7

Seq. No. 8869

Contig ID 14526_1.R1040

5'-most EST LIB3170-013-Q1-J1-E9

Seq. No. 8870

Contig ID 14526_2.R1040

5'-most EST LIB3170-019-Q1-J1-E10

Seq. No. 8871

Contig ID 14528_1.R1040

5'-most EST LIB3040-001-Q1-E1-D7

Seq. No. 8872

Contig ID 14534_2.R1040

5'-most EST LIB3040-001-Q1-E1-E2

Seq. No. 8873

Contig ID 14538_1.R1040 5'-most EST uC-gmropic064c03b1

Method BLASTX
NCBI GI g2492825
BLAST score 615
E value 2.0e-63
Match length 364
% identity 38

NCBI Description N-CARBAMYL-L-AMINO ACID AMIDOHYDROLASE

>gi 538915 pir JN0885 N-carbamyl-L-amino acid

amidohydrolase (EC 3.5.-.-) - Bacillus stearothermophilus

(strain NS1122A) >gi_460895_bbs_141971 (S67784) N-carbamyl-L-amino acid amidohydrolase [Bacillus

stearothermophilus, NS1122A, Peptide, 409 aa] [Bacillus

stearothermophilus]

Seq. No. 8874

Contig ID 14538_2.R1040 5'-most EST asn701134260.h1

Method BLASTX NCBI GI g3249039



BLAST score 250 E value 4.0e-21 Match length 132 % identity 42

NCBI Description (AF071221) N-carbamyl-L-amino acid amidohydrolase

[Arthrobacter aurescens]

Seq. No. 8875

Contig ID 14539_1.R1040 5'-most EST kmv700739010.h1

Method BLASTX
NCBI GI g4531443
BLAST score 954
E value 1.0e-103
Match length 264
% identity 73

NCBI Description (AC006224) putative 50s ribosomal protein L3 [Arabidopsis

thaliana]

Seq. No. 8876

Contig ID 14539 2.R1040

5'-most EST LIB3049-051-Q1-E1-H5

Method BLASTX
NCBI GI g4531443
BLAST score 383
E value 5.0e-37
Match length 127
% identity 65

NCBI Description (AC006224) putative 50s ribosomal protein L3 [Arabidopsis

thaliana]

Seq. No. 8877

Contig ID 14540 1.R1040

5'-most EST LIB3049-047-Q1-E1-B2

Seq. No. 8878

Contig ID 14540_2.R1040 5'-most EST leu701157315.h1

Seq. No. 8879

Contig ID 14553 1.R1040

5'-most EST LIB3073-004-Q1-K1-H7

Method BLASTN
NCBI GI g11570
BLAST score 139
E value 3.0e-72
Match length 155
% identity 97

NCBI Description Soybean chloroplast DNA for ribosomal protein L20

Seq. No. 8880

Contig ID 14586 1.R1040

5'-most EST LIB3170-078-Q1-J1-A2

Seq. No. 8881

Contig ID 14605 1.R1040

5'-most EST LIB3040-024-Q1-E1-C1



```
BLASTN
Method
                  g1139584
NCBI GI
BLAST score
                  189
                  1.0e-102
E value
Match length
                  341
                  89
% identity
                  Lupinus albus tRNA nucleotidyltransferase (CCA1) mRNA,
NCBI Description
                  complete cds
                   8882
Seq. No.
                   14608 1.R1040
Contig ID
                  LIB3072-018-Q1-E1-B1
5'-most EST
                   BLASTX
Method
                   q3551257
NCBI GI
BLAST score
                   536
                   1.0e-54
E value
                   178
Match length
% identity
                   (AB012708) 98b [Daucus carota]
NCBI Description
                   8883
Seq. No.
                   14608 2.R1040
Contig ID
                   leu701146919.h1
5'-most EST
Method
                   BLASTX
                   g3551257
NCBI GI
BLAST score
                   438
                   2.0e-43
E value
                   107
Match length
                   29
% identity
                   (AB012708) 98b [Daucus carota]
NCBI Description
                   8884
Seq. No.
                   14615 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220125f02a1
                   8885
Seq. No.
                   14622 1.R1040
Contig ID
                   LIB3040-028-Q1-E1-H3
5'-most EST
                   8886
 Seq. No.
                   14626 1.R1040
 Contig ID
                   LIB3040-029-Q1-E1-G9
 5'-most EST
                   8887
 Seq. No.
                    14629_1.R1040
 Contig ID
                   uC-gmronoir060d08b1
 5'-most EST
                   BLASTX
 Method
                   q4467130
 NCBI GI
 BLAST score
                    383
                    1.0e-36
 E value
                    130
 Match length
                    61
 % identity
                    (AL035538) glycosyltransferase like protein [Arabidopsis
```

thaliana] Seq. No. 8888

NCBI Description

Contig ID

14647 1.R1040



```
5'-most EST
                   LIB3040-038-Q1-E1-H2
                  8889
Seq. No.
                  14671 1.R1040
Contig ID
5'-most EST
                  LIB3170-015-Q1-J1-D1
                   8890
Seq. No.
                   14675 1.R1040
Contig ID
5'-most EST
                  LIB3040-050-Q1-E1-D12
                   8891
Seq. No.
Contig ID
                   14692 1.R1040
5'-most EST
                  LIB3139-061-P1-N1-F11
Seq. No.
                   8892
                   14692 2.R1040
Contig ID
5'-most EST
                   uC-qmflminsoy029f10b1
Seq. No.
                   8893
                   14696 1.R1040
Contig ID
                   leu701156412.h1
5'-most EST
Method
                   BLASTN
                   g11576
NCBI GI
BLAST score
                   634
E value
                   0.0e + 00
                   689
Match length
                   98
% identity
                  Soybean plastid DNA for rps12, rps7, 16S rRNA, tRNA-Val,
NCBI Description
                   NADH dehydrogenase and ORF
                   8894
Seq. No.
Contig ID
                   14696 3.R1040
5'-most EST
                   hyd700726585.h1
Method
                   BLASTN
NCBI GI
                   g11576
BLAST score
                   83
E value
                   5.0e-39
                   209
Match length
                   99
% identity
NCBI Description
                  Soybean plastid DNA for rps12, rps7, 16S rRNA, tRNA-Val,
                   NADH dehydrogenase and ORF
                   8895
Seq. No.
Contig ID
                   14702 1.R1040
5'-most EST
                   LIB3170-018-Q1-K1-C11
Seq. No.
                   8896
Contig ID
                   14708 1.R1040
5'-most EST
                   jC-gmst02400029d10a1
```

Method BLASTX
NCBI GI g3540204
BLAST score 328
E value 2.0e-30
Match length 114
% identity 52

NCBI Description (AC004260) Hypothetical protein [Arabidopsis thaliana]



Seq. No. 8897

Contig ID 14708 2.R1040

5'-most EST LIB3138-065-Q1-N1-G9

Method BLASTX
NCBI GI g3540204
BLAST score 312
E value 9.0e-29
Match length 76
% identity 71

NCBI Description (AC004260) Hypothetical protein [Arabidopsis thaliana]

Seq. No. 8898

Contig ID 14710 1.R1040

5'-most EST jC-gmro02910005f04d1

Seq. No. 8899

Contig ID 14711 1.R1040

5'-most EST LIB3049-056-Q1-E1-H1

Method BLASTX
NCBI GI g2498490
BLAST score 458
E value 3.0e-93
Match length 320
% identity 54

NCBI Description VIRAL INTEGRATION SITE PROTEIN INT-6 >qi 1854579 (L35556)

Int-6 [Mus musculus] >gi_2114363 (U62962) similar to mouse
Int-6 [Homo sapiens] >gi_2351382 (U54562) eIF3-p48 [Homo

sapiens] >gi 2688818 (U85947) Int-6 [Homo sapiens]

>gi_2695701 (U94175) mammary tumor-associated protein INT6 [Homo sapiens] >gi_4503521_ref_NP_001559.1_pEIF3S6_ murine

mammary tumor integration site 6 (oncogene homolog)

Seq. No. 8900

Contig ID 14713_1.R1040 5'-most EST uC-gmropic004e10b1

Method BLASTX
NCBI GI g2815493
BLAST score 1296
E value 1.0e-143
Match length 376
% identity 62

NCBI Description SERINE CARBOXYPEPTIDASE I PRECURSOR (CARBOXYPEPTIDASE C)

(CP-MI) >gi_1731988_emb_CAA70816_ (Y09603) serine

carboxypeptidase I, CP-MI [Hordeum vulgare]

Seq. No. 8901

Contig ID 14713 2.R1040

5'-most EST LIB3049-056-Q1-E1-H11

Method BLASTX
NCBI GI g2815493
BLAST score 666
E value 8.0e-70
Match length 174
% identity 66

NCBI Description SERINE CARBOXYPEPTIDASE I PRECURSOR (CARBOXYPEPTIDASE C)

(CP-MI) >gi_1731988_emb_CAA70816_ (Y09603) serine

carboxypeptidase I, CP-MI [Hordeum vulgare]



```
Seq. No. 8902
```

Contig ID 14714 1.R1040

5'-most EST LIB3049-056-Q1-E1-H12

Method BLASTX
NCBI GI g2673908
BLAST score 202
E value 1.0e-15
Match length 110
% identity 41

NCBI Description (AC002561) hypothetical protein [Arabidopsis thaliana]

Seq. No.

8903

Contig ID 5'-most EST

14714_2.R1040 fua701043419.h1

Method BLASTX
NCBI GI g2673908
BLAST score 373
E value 2.0e-35
Match length 217
% identity 47

NCBI Description (AC002561) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8904

Contig ID 14717_1.R1040 5'-most EST asn701131988.h1

Method BLASTX
NCBI GI g2244920
BLAST score 581
E value 7.0e-60
Match length 177
% identity 62

NCBI Description (Z97339) hypothetical protein [Arabidopsis thaliana]

Seq. No.

8905

Contig ID

14731 1.R1040

5'-most EST

LIB3049-056-Q1-E1-F7

Seq. No.

8906

Contig ID 5'-most EST

14734_1.R1040 leu701146805.h1

Method BLASTX
NCBI GI g1711512
BLAST score 537
E value 6.0e-55
Match length 127
% identity 76

NCBI Description SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 2 (SRP54)

>gi_1076577_pir__S51598 signal recognition particle 54K
protein - tomato (cv. UC82-B) >gi_556902_emb_CAA84288_
(Z34527) 54-kD signal recognition particle (SRP) specific

protein [Lycopersicon esculentum]

Seq. No.

8907

Contig ID 5'-most EST

14734_2.R1040 zhf700953055.h1

Method BLASTX



NCBI GI g1711512 BLAST score 443 E value 5.0e-44 Match length 97 % identity 92

NCBI Description SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 2 (SRP54)

>gi_1076577_pir__S51598 signal recognition particle 54K

protein - tomato (cv. UC82-B) >gi_556902_emb_CAA84288_ (Z34527) 54-kD signal recognition particle (SRP) specific

protein [Lycopersicon esculentum]

Seq. No. 8908

Contig ID 14739 1.R1040

5'-most EST LIB3049-056-Q1-E1-E9

Method BLASTX
NCBI GI g3980377
BLAST score 302
E value 3.0e-27
Match length 83
% identity 69

NCBI Description (AC004561) unknown protein [Arabidopsis thaliana]

Seq. No. 8909

Contig ID 14746_1.R1040 5'-most EST seb700651468.h1

Method BLASTX
NCBI GI g3128186
BLAST score 347
E value 2.0e-32
Match length 183
% identity 52

NCBI Description (AC004521) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8910

Contig ID 14746 2.R1040

5'-most EST LIB3138-029-Q1-N1-H5

Seq. No. 8911

Contig ID 14747_1.R1040

5'-most EST LIB3049-056-Q1-E1-D9

Seq. No. 8912

Contig ID 14755_1.R1040 5'-most EST fde700876676.h1

Method BLASTX
NCBI GI g2827536
BLAST score 192
E value 1.0e-14
Match length 60
% identity 53

NCBI Description (AL021633) hypothetical protein [Arabidopsis thaliana]

Seq. No. 8913

Contig ID 14762 1.R1040

5'-most EST LIB3170-039-01-K2-H3

Seq. No. 8914



```
14762 2.R1040
Contig ID
5'-most EST
                  LIB3109-002-Q1-K3-G1
Seq. No.
                  8915
Contig ID
                  14778 1.R1040
5'-most EST
                  LIB3049-056-Q1-E1-A11
Seq. No.
Contig ID
                  14779 1.R1040
5'-most EST
                  LIB3049-056-Q1-E1-A12
Method
                  BLASTX
NCBI GI
                  q4531444
BLAST score
                  1632
```

E value 0.0e + 00Match length 465 % identity 66

NCBI Description (AC006224) putative protein kinase [Arabidopsis thaliana]

Seq. No. 8917 14779 2.R1040 Contig ID 5'-most EST dpv701097679.h1 BLASTX Method NCBI GI q4531444

BLAST score 261 E value 2.0e-22 Match length 62 % identity

NCBI Description (AC006224) putative protein kinase [Arabidopsis thaliana]

Seq. No. 8918

Contig ID 14780 1.R1040

5'-most EST LIB3049-056-Q1-E1-A2

8919 Seq. No.

Contig ID 14780 2.R1040 5'-most EST dkc700968064.h1

Seq. No. 8920

Contig ID 14782 1.R1040

5'-most EST LIB3049-056-Q1-E1-A4

Method BLASTX NCBI GI q4432835 BLAST score 289 E value 7.0e-26 Match length 88 % identity 57

NCBI Description (AC006283) unknown protein [Arabidopsis thaliana]

8921 Seq. No.

14782 2.R1040 Contig ID

 $LIB30\overline{4}9-042-Q1-E1-A4$ 5'-most EST

Method BLASTX NCBI GI g4432835 BLAST score 241 2.0e-20 E value Match length 76 % identity 54

1599

A super section



NCBI Description (AC006283) unknown protein [Arabidopsis thaliana]

Seq. No. 8922

Contig ID 14785 1.R1040

5'-most EST LIB3109-011-Q1-K1-A9

Method BLASTX
NCBI GI g4539452
BLAST score 1172
E value 1.0e-129
Match length 289
% identity 71

NCBI Description (AL049500) putative phosphoribosylanthranilate transferase

[Arabidopsis thaliana]

Seq. No. 8923

Contig ID 14795_1.R1040 5'-most EST asn701139392.h1

Method BLASTN
NCBI GI 9455949
BLAST score 39
E value 2.0e-12
Match length 39
% identity 100

NCBI Description 18S rRNA {3' region} [Glycine max=soybeans, var. Pella 86,

leaves, rRNA Partial, 111 nt]

Seq. No. 8924

Contig ID 14795 2.R1040

5'-most EST LIB3073-001-Q1-K1-H4

Seq. No. 8925

Contig ID 14796_1.R1040 5'-most EST awf700837919.h1

Method BLASTX
NCBI GI g100226
BLAST score 167
E value 2.0e-11
Match length 76
% identity 43

NCBI Description hypothetical protein - tomato >gi_19275_emb_CAA78112_

(Z12127) protein of unknown function [Lycopersicon esculentum] >gi_445619 prf_ 1909366A Leu zipper protein

[Lycopersicon esculentum]

Seq. No. 8926

Contig ID 14796_3.R1040 5'-most EST fde700874185.h1

Seq. No. 8927

Contig ID 14800 1.R1040 5'-most EST zzp700830261.h1

Seq. No. 8928

Contig ID 14807 1.R1040

5'-most EST LIB3170-038-Q1-K1-G2

Method BLASTX NCBI GI g3033377



```
BLAST score
                    103
                    1.0e-09
 E value
 Match length
                    119
 % identity
                    34
 NCBI Description (AC004238) putative berberine bridge enzyme [Arabidopsis
                    thaliana]
 Seg. No.
                    8929
 Contig ID
                    14808 1.R1040
 5'-most EST
                    LIB3170-039-Q1-K2-H2
 Method
                    BLASTX
 NCBI GI
                    g4544386
 BLAST score
                    685
                    3.0e-72
 E value
 Match length
                    157
                    79
 % identity
 NCBI Description (AC007047) putative cell division control protein
                    [Arabidopsis thaliana]
                    8930
 Seq. No.
                    14809 2.R1040
 Contig ID
 5'-most EST
                    zhf700961825.h1
 Method
                    BLASTX
 NCBI GI
                    q3047104
 BLAST score
                    565
 E value
                    3.0e-58
 Match length
                    157
                    70
 % identity
 NCBI Description (AF058919) No definition line found [Arabidopsis thaliana]
 Seq. No.
                    8931
                    14810 1.R1040
 Contig ID
 5'-most EST
                    LIB3049-055-Q1-E1-E5
 Seq. No.
                    8932
 Contig ID
                    14818 1.R1040
 5'-most EST
                    LIB3049-055-Q1-E1-F4
                    8933
 Seq. No.
 Contig ID
                    14820 1.R1040
 5'-most EST
                    ssr700555180.hl
 Seq. No.
                    8934
                    14821 1.R1040
 Contig ID
 5'-most EST
                    LIB3049-055-Q1-E1-F7
 Method
                    BLASTX
 NCBI GI
                    g3915463
 BLAST score
                    225
 E value
                    5.0e-23
 Match length
                    136
 % identity
                    46
                    HYPOTHETICAL 33.1 KD PROTEIN SLR1592
 NCBI Description
```

>gi_1652063 dbj_BAA16988 (D90902) hypothetical protein

[Synechocystis sp.]

Seq. No. 8935

Contig ID 14822 1.R1040



```
5'-most EST
                  LIB3170-040-Q1-K1-B3
                  8936
Seq. No.
Contig ID
                  14824 1.R1040
                  kl1701209491.h1
5'-most EST
Method
                  BLASTX
                  g2598575
NCBI GI
                  293
BLAST score
E value
                  3.0e-26
Match length
                  120
% identity
                  47
NCBI Description
                 (Y15293) MtN21 [Medicago truncatula]
                  8937
Seq. No.
                  14829 1.R1040
Contig ID
5'-most EST
                  leu701148315.hl
                  8938
Seq. No.
                  14831 1.R1040
Contig ID
5'-most EST
                  LIB3051-036-Q1-K1-G8
Method
                  BLASTX
NCBI GI
                  g2832700
                  229
BLAST score
E value
                  2.0e-47
                  152
Match length
% identity
                  68
NCBI Description (AL021713) unknown protein [Arabidopsis thaliana]
Seq. No.
                  8939
Contig ID
                  14832 1.R1040
5'-most EST
                  LIB3049-012-Q1-E1-D1
Method
                  BLASTX
NCBI GI
                  q3297808
BLAST score
                  248
E value
                  4.0e-21
Match length
                  97
% identity
                  52
NCBI Description (AL031032) putative protein [Arabidopsis thaliana]
Seq. No.
                  8940
                  14834 1.R1040
Contig ID
5'-most EST
                  LIB3049-055-Q1-E1-D6
                  8941
Seq. No.
Contig ID
                  14847 1.R1040
5'-most EST
                  LIB3049-055-Q1-E1-B3
                  8942
Seq. No.
Contig ID
                  14848 1.R1040
5'-most EST
                  kl1701211013.h1
Seq. No.
                  8943
                  14850 1.R1040
Contig ID
5'-most EST
                  LIB3170-037-Q1-K1-H2
```

1602

8944

14851 1.R1040

Seq. No.

Contig ID



```
5'-most EST
                   LIB3170-039-Q1-K2-A2
Method
                   BLASTX
NCBI GI
                   g3860263
BLAST score
                   358
E value
                   9.0e-34
Match length
                   111
% identity
                   59
NCBI Description
                   (AC005824) putative cytochrome p450 protein [Arabidopsis
                   thaliana]
                   8945
Seq. No.
                   14853 1.R1040
Contig ID
5'-most EST
                   LIB3049-055-Q1-E1-B9
Method
                   BLASTX
NCBI GI
                   g2208946
BLAST score
                   451
E value
                   4.0e-45
Match length
                   102
% identity
NCBI Description
                  (Y10116) signal recognition particle subunit 14
                   [Arabidopsis thaliana]
                   8946
Seq. No.
                  14854 1.R1040
Contig ID
                  LIB3170-023-Q1-K1-F7
5'-most EST
Method
                  BLASTX
NCBI GI
                   g4508073
BLAST score
                   411
E value
                   3.0e-40
Match length
                  123
% identity
                   65
NCBI Description
                  (AC005882) 43220 [Arabidopsis thaliana]
Seq. No.
                   8947
                   14860 1.R1040
Contig ID
5'-most EST
                  LIB3049-054-Q1-E1-H3
                   8948
Seq. No.
                   14862 1.R1040
Contig ID
5'-most EST
                  g5058294
Seq. No.
                   8949
                   14863 1.R1040
Contig ID
                  LIB3170-009-Q2-J1-E12
5'-most EST
Method
                  BLASTX
NCBI GI
                   q4220462
BLAST score
                   526
```

E value 1.0e-53 Match length 110 % identity 86

NCBI Description (AC006216) Strong similarity to gb_Z50851 HD-zip (athb-8)

gene from Arabidopsis thaliana containing Homeobox PF 00046

and bZIP PF_00170 domains. [Arabidopsis thaliana]

Seq. No. 8950

14865 1.R1040 Contig ID

5'-most EST LIB3049-051-Q1-E1-E9



Method BLASTX
NCBI GI g4455800
BLAST score 404
E value 8.0e-39
Match length 145
% identity 55

NCBI Description (Z97343) unnamed protein product [Arabidopsis thaliana]

Seq. No. 8951

Contig ID 14865 2.R1040

5'-most EST jC-gmro02800031g07a1

Method BLASTX
NCBI GI g1729891
BLAST score 184
E value 1.0e-13
Match length 108
% identity 35

NCBI Description TEGT PROTEIN (TESTIS ENHANCED GENE TRANSCRIPT)

>gi_2136254_pir__I38334 TEGT (testis enhanced gene

transcript) - human >gi_458545 emb_CAA53472 (X75861) TEGT [Homo sapiens] >gi_4507433 ref_NP_003208.1_pTEGT_testis

enhanced gene transcript

Seq. No. 8952

Contig ID 14869 1.R1040

5'-most EST LIB3049-055-Q1-E1-A11

Seq. No. 8953

Contig ID 14869 2.R1040

5'-most EST LIB3109-018-Q1-K1-A8

Method BLASTX
NCBI GI g3047102
BLAST score 187
E value 5.0e-14
Match length 86
% identity 43

NCBI Description (AF058919) No definition line found [Arabidopsis thaliana]

Seq. No. 8954

Contig ID 14872_1.R1040 5'-most EST g5126928

Seq. No.

8955

Contig ID 14873 1.R1040

5'-most EST jC-gmf102220097c01a1

Method BLASTN
NCBI GI g2275194
BLAST score 226
E value 1.0e-124
Match length 662
% identity 84

NCBI Description Arabidopsis thaliana chromosome II BAC T08I13 genomic

sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 8956

Contig ID 14882_1.R1040

5'-most EST LIB3051-058-Q1-K2-F4

Method



BLASTX

```
g2708624
NCBI GI
                   1543
BLAST score
                   1.0e-172
E value
                   358
Match length
                   76
% identity
NCBI Description (AF036618) acetyl-CoA synthetase [Arabidopsis thaliana]
Seq. No.
                   8957
                   14888 1.R1040
Contig ID
                   LIB3170-039-Q1-K2-E1
5'-most EST
                   BLASTX
Method
                   q2160173
NCBI GI
BLAST score
                   179
                   5.0e-13
E value
                   78
Match length
% identity
                    (AC000132) Similar to N. tabacum salt-inducible protein
NCBI Description
                    (gb U08285). [Arabidopsis thaliana]
                   8958
Seq. No.
                   14891 1.R1040
Contig ID
                    jC-gmst02400073a05a1
5'-most EST
Method
                   BLASTX
                    q4432823
NCBI GI
BLAST score
                    433
                    2.0e-42
E value
Match length
                    240
                    45
% identity
                    (AC006593) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                    8959
Seq. No.
                    14910 1.R1040
Contig ID
                    k1170\overline{1}205175.h1
 5'-most EST
Method
                    BLASTX
                    g1172874
NCBI GI
                    210
BLAST score
                    2.0e-16
 E value
                    149
Match length
                    43
 % identity
                    DEHYDRATION-RESPONSIVE PROTEIN RD22 PRECURSOR
 NCBI Description
                    >gi_479589_pir__S34823 dehydration-induced protein RD22 -
                    Arabidopsis thaliana >gi_391608_dbj_BAA01546_ (D10703) rd22
                    [Arabidopsis thaliana] > \overline{gi}_4471\overline{3}4_prf__19134\overline{2}1A rd22 gene
                    [Arabidopsis thaliana]
                    8960
 Seq. No.
                    14912 2.R1040
 Contig ID
 5'-most EST
                    LIB3049-054-Q1-E1-B7
                    BLASTX
 Method
                    q2252838
 NCBI GI
                    159
 BLAST score
                    2.0e-10
 E value
 Match length
                    155
 % identity
                    27
                   (AF013293) No definition line found [Arabidopsis thaliana]
 NCBI Description
```



56

% identity

NCBI Description

```
Seq. No.
                  8961
Contig ID
                  14915 1.R1040
5'-most EST
                  gsv701055256.hl
Method
                  BLASTX
NCBI GI
                  g2244850
                  257
BLAST score
E value
                  5.0e-22
Match length
                  111
                  52
% identity
NCBI Description
                 (Z97337) hypothetical protein [Arabidopsis thaliana]
                  8962
Seq. No.
                  14915 2.R1040
Contig ID
                  kl1701212335.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2244850
BLAST score
                  188
                   4.0e-14
E value
Match length
                  85
% identity
                   51
NCBI Description (Z97337) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   8963
                  14917 1.R1040
Contig ID
5'-most EST
                  LIB3049-054-Q1-E1-C12
Method
                  BLASTN
NCBI GI
                  g3449331
BLAST score
                  33
                  8.0e-09
E value
                  73
Match length
% identity
                  86
NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
                  MNC17, complete sequence [Arabidopsis thaliana]
Seq. No.
                   8964
Contig ID
                   14924 1.R1040
5'-most EST
                   fde700877065.hl
Method
                  BLASTX
NCBI GI
                  g4191779
BLAST score
                   404
E value
                   1.0e-39
Match length
                  100
                  74
% identity
NCBI Description
                  (AC005917) putative recA protein [Arabidopsis thaliana]
                   8965
Seq. No.
                  14925_1.R1040
Contig ID
                  LIB3049-008-Q1-E1-G12
5'-most EST
Method
                  BLASTX
NCBI GI
                   q2618723
BLAST score
                   387
                   6.0e-37
E value
                  173
Match length
```

1606

(U49073) IAA17 [Arabidopsis thaliana] >gi 2921756

(AF040631) IAA17/AXR3 protein [Arabidopsis thaliana] >gi_4389514_gb_AAB70451 (AC000104) Identical to



Arabidopsis qb AF040632 and gb U49073 IAA17/AXR3 gene. ESTs gb H36782 and gb F14074 come from this gene. [Arabidopsis thaliana]

Seq. No. Contig ID 5'-most EST 8966

14926 1.R1040 ssr700555756.hl

Seq. No.

8967

Contig ID

14926 2.R1040

5'-most EST

uC-gmflminsoy078f07b1

Seq. No.

8968

Contig ID

14926 3.R1040

5'-most EST

LIB3170-035-Q1-K1-F10

Seq. No.

8969

Contig ID

14931 1.R1040

5'-most EST

LIB3049-046-Q1-E1-H4

Seq. No.

8970

Contig ID

14932 1.R1040

5'-most EST

LIB3170-033-Q1-K1-F12

Seq. No.

8971

Contig ID

14945 1.R1040

5'-most EST

LIB3170-034-Q1-K1-H2

Method

BLASTN

NCBI GI

g1420886

BLAST score

124

E value

5.0e-63

Match length

302

% identity

NCBI Description Phaseolus vulgaris non-specific lipid transfer-like protein

mRNA, complete cds

Seq. No.

8972

Contig ID

14947 1.R1040

5'-most EST Method

LIB3093-038-Q1-K1-A6 BLASTX

NCBI GI

g1352267

405

BLAST score

2.0e-39

E value

147

Match length % identity

NCBI Description

DEOXYHYPUSINE SYNTHASE >gi 994715 (L39068) deoxyhypusine synthase [Homo sapiens] >gi_1710220 (U79262) deoxyhypusine

synthase [Homo sapiens] >gi_3021398_emb_CAA04940

(AJ001701) deoxyhypusine synthase [Homo sapiens]

>gi 4503325 ref NP 001921.1 pDHPS deoxyhypusine synthase

Seq. No.

8973

Contig ID

14948 1.R1040

5'-most EST

trc700567438.h1

Seq. No.

8974

Contig ID

14949 1.R1040



5'-most EST LIB3049-010-Q1-E1-H5 Method BLASTX NCBI GI q3643609 BLAST score 541 5.0e-55 E value Match length 158 70 % identity NCBI Description (AC005395) putative Cys3His zinc finger protein [Arabidopsis thaliana] 8975 Seq. No. 14949 2.R1040 Contig ID 5'-most EST g4395984 Method BLASTX NCBI GI q3687237 BLAST score 163 E value 4.0e-11 Match length 57 67 % identity NCBI Description (AC005169) putative Cys3His zinc-finger protein [Arabidopsis thaliana] Seq. No. 8976 14951 1.R1040 Contig ID 5'-most EST LIB3049-053-Q1-E1-G10 Seq. No. 8977 Contig ID 14953 1.R1040 5'-most EST LIB3170-034-Q1-K1-H10 Seq. No. 8978 Contig ID 14955 1.R1040 5'-most EST LIB3049-053-Q1-E1-G3 Method BLASTX NCBI GI g2464914 BLAST score 501 1.0e-50 E value Match length 146 70 % identity NCBI Description (Z99708) hypothetical protein [Arabidopsis thaliana] 8979 Seq. No. 14955 2.R1040 Contig ID 5'-most EST LIB3107-033-Q1-K1-G12 Method BLASTX NCBI GI q2464914 BLAST score 195 E value 5.0e-15 Match length 71 % identity 61 (Z99708) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 8980

Contig ID 14960_1.R1040

5'-most EST LIB3139-007-P1-N1-A7

Method BLASTX NCBI GI g1488043



```
1846
BLAST score
                  0.0e + 00
E value
                  438
Match length
                  80
% identity
NCBI Description (U63784) PAPS-reductase-like protein [Catharanthus roseus]
                  8981
Seq. No.
                  14960 2.R1040
Contig ID
                  gsv701047450.hl
5'-most EST
                  BLASTX
Method
```

q1488043 NCBI GI 210 BLAST score 7.0e-17 E value 100 Match length 49 % identity

(U63784) PAPS-reductase-like protein [Catharanthus roseus] NCBI Description

Seq. No.

14960 3.R1040 Contig ID kl1701202445.h1 5'-most EST

8983 Seq. No.

14963 1.R1040 Contig ID $seb70\overline{0}653658.h1$ 5'-most EST

Method BLASTX q731520 NCBI GI 196 BLAST score 8.0e-15 E value 80 Match length 46 % identity

CYTOCHROME C OXIDASE ASSEMBLY PROTEIN COX15 NCBI Description

>gi_1077711_pir__S50644 cytochrome oxidase assembly factor

COX15 - yeast (Saccharomyces cerevisiae) >gi 603381 (U18917) Cox15p: Cytochrom oxidase assembly factor

[Saccharomyces cerevisiae] >gi_603947 (L38643) cytochrome

oxidase assembly factor [Saccharomyces cerevisiae]

8984 Seq. No.

14963 2.R1040 Contig ID fC-gmse700754560a2 5'-most EST

8985 Seq. No.

14964 1.R1040 Contig ID

LIB3051-009-Q1-E1-C9 5'-most EST

Method BLASTX q2245030 NCBI GI BLAST score 306 3.0e-27 E value 207 Match length 40

% identity (Z97341) apetala2 domain TINY homolog [Arabidopsis NCBI Description

thaliana]

8986 Seq. No.

Contig ID 14965 1.R1040

LIB3049-053-Q1-E1-E1 5'-most EST



Seq. No. 8987

Contig ID 14970 1.R1040

5'-most EST LIB3170-033-Q1-J1-E12

Seq. No. 8988

Contig ID 14971 1.R1040 5'-most EST uaw700666538.h1

Method BLASTX
NCBI GI g3776578
BLAST score 206
E value 3.0e-16
Match length 80
% identity 49

NCBI Description (AC005388) ESTs gb F13915 and gb F13916 come from this

gene. [Arabidopsis thaliana]

Seq. No. 8989

Contig ID 14975_1.R1040 5'-most EST awf700838186.h1

Seq. No. 8990

Contig ID 14977_1.R1040 5'-most EST rlr700901579.h1

Method BLASTX
NCBI GI g2459431
BLAST score 164
E value 6.0e-11
Match length 97
% identity 47

NCBI Description (AC002332) unknown protein [Arabidopsis thaliana]

Seq. No. 8991

Contig ID 14979_1.R1040 5'-most EST fde700874980.h1

Seq. No. 8992

Contig ID 14982 1.R1040

5'-most EST LIB3049-053-Q1-E1-B9

Seq. No. 8993

Contig ID 14986_1.R1040

5'-most EST LIB3074-029-Q1-K1-D5

Seq. No. 8994

Contig ID 14988_1.R1040

5'-most EST LIB $30\overline{4}$ 9-053-Q1-E1-C4

Method BLASTX
NCBI GI g3688171
BLAST score 373
E value 1.0e-35
Match length 168
% identity 55

NCBI Description (AL031804) putative protein [Arabidopsis thaliana]

Seq. No. 8995

Contig ID 14992_1.R1040

5'-most EST LIB3093-045-Q1-K1-E7



Method BLASTX
NCBI GI g3150406
BLAST score 685
E value 6.0e-72
Match length 317
% identity 44

NCBI Description (AC004165) putative indole-3-acetate

beta-glucosyltransferase [Arabidopsis thaliana]

Seq. No. 8996

Contig ID 14993 1.R1040

5'-most EST LIB3049-053-Q1-E1-D11

Method BLASTX
NCBI GI g4039152
BLAST score 182
E value 3.0e-13
Match length 53
% identity 64

NCBI Description (AF104221) low temperature and salt responsive protein

LTI6B [Arabidopsis thaliana] >gi_4325219_gb_AAD17303_(AF122006) hydrophobic protein [Arabidopsis thaliana]

Seq. No. 8997

Contig ID 14993_2.R1040 5'-most EST taw700657926.h1

Method BLASTX
NCBI GI g4039152
BLAST score 182
E value 3.0e-13
Match length 53
% identity 64

NCBI Description (AF104221) low temperature and salt responsive protein

LTI6B [Arabidopsis thaliana] >gi_4325219_gb_AAD17303_(AF122006) hydrophobic protein [Arabidopsis thaliana]

Seq. No. 8998

Contig ID 14997_1.R1040 5'-most EST txt700731851.h1

Seq. No. 8999

Contig ID 14997_2.R1040 5'-most EST pxt700945886.h1

Seq. No. 9000

Contig ID 14998_1.R1040 5'-most EST eep700866824.h1

Method BLASTX
NCBI GI g4160280
BLAST score 245
E value 2.0e-20
Match length 165
% identity 42

NCBI Description (AJ006224) purple acid phosphatase [Ipomoea batatas]

Seq. No. 9001

Contig ID 14999_1.R1040

5'-most EST jC-gmfl02220141d07a1



```
BLASTX
Method
                   q3785997
NCBI GI
                   622
BLAST score
                   9.0e-65
E value
                   200
Match length
                   62
% identity
                   (AC005499) putative annexin [Arabidopsis thaliana]
NCBI Description
                   9002
Seq. No.
                   15000 1.R1040
Contig ID
                   uC-gmronoir053e10b1
5'-most EST
                   9003
Seq. No.
                   15000 2.R1040
Contig ID
                   dpv701103624.h1
5'-most EST
                   9004
Seq. No.
                   15002 1.R1040
Contig ID
                   uC-gmflminsoy091h05b1
5'-most EST
                   BLASTN
Method
                   q3746902
NCBI GI
                   286
BLAST score
                   1.0e-159
E value
                   517
Match length
                    89
% identity
NCBI Description
```

Pisum sativum signal recognition particle 54 kDa subunit precursor (Ffc) mRNA, nuclear gene encoding chloroplast

protein, partial cds

9005 Seq. No.

15004 1.R1040 Contig ID

LIB $30\overline{4}9-053-Q1-E1-A6$ 5'-most EST

9006 Seq. No.

15007 1.R1040 Contig ID

LIB3049-040-Q1-E1-G11 5'-most EST

Method BLASTX q1076387 NCBI GI 789 BLAST score 3.0e-84E value 181 Match length 83 % identity

protein kinase homolog - Arabidopsis thaliana NCBI Description

>gi_717180_emb_CAA55866_ (X79279) protein kinase homologous

to shaggy and glycogen synthase kinase-3 [Arabidopsis

thaliana]

9007 Seq. No.

15007 2.R1040 Contig ID fC-qmle700872092d2 5'-most EST

BLASTN Method g1504062 NCBI GI 241 BLAST score 1.0e-132 E value 703 Match length 84 % identity

NCBI Description A.thaliana mRNA for shaggy-like kinase kappa



9008 Seq. No.

15007 3.R1040 Contig ID

LIB3049-053-Q1-E1-A9 5'-most EST

BLASTX Method NCBI GI q1076387 BLAST score 187 2.0e-18 E value Match length 70 % identity

NCBI Description

protein kinase homolog - Arabidopsis thaliana >gi_717180_emb_CAA55866_ (X79279) protein kinase homologous

to shaggy and glycogen synthase kinase-3 [Arabidopsis

thaliana]

9009 Seq. No.

15007 4.R1040 Contig ID

5'-most EST LIB3109-012-Q1-K1-E6

Seq. No.

15008 1.R1040 Contig ID 5'-most EST $seb70\overline{0}652702.h1$

Seq. No.

15008 2.R1040 Contig ID 5'-most EST $seb70\overline{0}652769.h1$

9012 Seq. No.

15008 3.R1040 Contig ID epx701106574.h1 5'-most EST

9013 Seq. No.

15013 1.R1040 Contig ID

iC-qmst02400009b04a1 5'-most EST

Seq. No. 9014

15023 1.R1040 Contig ID $uaw70\overline{0}664606.h1$ 5'-most EST

BLASTX Method g1255951 NCBI GI 1406 BLAST score 1.0e-156 E value 363 Match length 75 % identity

(X96932) PS60 [Nicotiana tabacum] NCBI Description

9015 Seq. No.

15027 1.R1040 Contig ID

LIB3138-071-P1-N1-B7 5'-most EST

BLASTX Method g3892056 NCBI GI BLAST score 568 3.0e-58 E value Match length 126 % identity 87

NCBI Description (AC002330) putative vacuolar ATPase [Arabidopsis thaliana]



```
9016
Seq. No.
                   15033 1.R1040
Contig ID
                   LIB31\overline{0}9-038-Q1-K1-D6
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2832435
BLAST score
                   205
                   3.0e-16
E value
                   102
Match length
% identity
                   42
                   (AJ223042) noisette [Drosophila melanogaster]
NCBI Description
                   9017
Seq. No.
                   15035 1.R1040
Contig ID
5'-most EST
                   LIB3074-025-Q1-E1-B9
Method
                   BLASTX
                   q4530585
NCBI GI
                   373
BLAST score
E value
                   1.0e-35
                   87
Match length
% identity
                   76
NCBI Description (AF130978) B12D protein [Ipomoea batatas]
Seq. No.
                   9018
                   15035 2.R1040
Contig ID
                   LIB3049-045-Q1-E1-D12
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4530585
                   276
BLAST score
                   3.0e-25
E value
Match length
                   78
% identity
                   76
                   (AF130978) B12D protein [Ipomoea batatas]
NCBI Description
                   9019
Seq. No.
                   15036 1.R1040
Contig ID
                   LIB3049-052-Q1-E1-E9
5'-most EST
                   9020
Seq. No.
                   15037 1.R1040
Contig ID
                   LIB3049-012-Q1-E1-C12
5'-most EST
Method
                   BLASTX
                   g3183088
NCBI GI
                   291
BLAST score
                    5.0e-26
E value
Match length
                   101
% identity
                    53
                   PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR
NCBI Description
                    (LTP) >gi 629658 pir S47084 lipid transfer like protein -
                    cowpea >g\bar{i} 49903\bar{4} emb CAA56113 (X79604) lipid transfer
                    like protein [Vigna unguiculata]
```

Seq. No. 9021

Contig ID 15037_2.R1040

5'-most EST LIB3049-047-Q1-E1-A5

Method BLASTX
NCBI GI g3183088
BLAST score 289



E value 6.0e-26

Match length 98 % identity 51

NCBI Description PROBABLE NONSPECIFIC LIPID-TRANSFER PROTEIN AKCS9 PRECURSOR

(LTP) >gi_629658_pir__S47084 lipid transfer like protein - cowpea >gi_499034_emb_CAA56113_ (X79604) lipid transfer

like protein [Vigna unguiculata]

Seq. No. 9022

Contig ID 15039_1.R1040 5'-most EST ncj700982409.h1

Method BLASTX
NCBI GI 93860257
BLAST score 266
E value 1.0e-22
Match length 136
% identity 43

NCBI Description (AC005824) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9023

Contig ID 15039_2.R1040 5'-most EST zzp700834617.h1

Seq. No. 9024

Contig ID 15040 1.R1040

5'-most EST LIB3072-046-Q1-K1-C12

Method BLASTX
NCBI GI g1085952
BLAST score 204
E value 4.0e-16
Match length 106
% identity 36

NCBI Description hypothetical protein S1 - Phalaris coerulescens

>gi_556831_emb_CAA57519_ (X81991) S1 [Phalaris
coerulescens] >gi_1103489_emb_CAA63108_ (X92351)

self-incompatibility mutant protein [Phalaris coerulescens]

Seq. No. 9025

Contig ID 15042 1.R1040

5'-most EST LIB3109-029-Q1-K1-C9

Method BLASTX
NCBI GI g3122287
BLAST score 146
E value 9.0e-09
Match length 152
% identity 24

NCBI Description PUTATIVE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE

(TRANSAMINASE B) (BCAT) >gi_2127785_pir__G64425

branched-chain amino acid aminotransferase - Methanococcus jannaschii >gi_1591667 (U67543) branched-chain amino acid

aminotransferase (ilvE) [Methanococcus jannaschii]

Seq. No. 9026

Contig ID 15043 1.R1040

5'-most EST jC-gmle01810067h04a1

Method BLASTX NCBI GI g1432058

NCBI Description



```
696
BLAST score
                   7.0e-82
E value
                   280
Match length
                   57
% identity
NCBI Description (U58540) WRKY2 [Petroselinum crispum]
                   9027
Seq. No.
                   15044 1.R1040
Contig ID
                   jC-qmst02400030c02a1
5'-most EST
                   BLASTX
Method
                   q2702272
NCBI GI
                   475
BLAST score
                   4.0e-56
E value
Match length
                   236
% identity
                   29
                   (AC003033) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   15046 1.R1040
Contig ID
                   uC-gmrominsoy271h03b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3878134
BLAST score
                   213
                   7.0e-17
E value
Match length
                   61
% identity
                   61
                   (Z68218) K01H12.1 [Caenorhabditis elegans]
NCBI Description
                    9029
Seq. No.
                   15046 2.R1040
Contig ID
                   LIB3049-052-Q1-E1-E11
5'-most EST
                   BLASTX
Method
                   g3878134
NCBI GI
                   213
BLAST score
                    4.0e-17
E value
                    61
Match length
                    59
 % identity
                   (Z68218) K01H12.1 [Caenorhabditis elegans]
NCBI Description
                    9030
Seq. No.
                    15051 1.R1040
Contig ID
                    sat701013119.h1
 5'-most EST
                    9031
 Seq. No.
                    15051 2.R1040
 Contig ID
                    LIB3072-002-Q1-K1-D12
 5'-most EST
                    9032
 Seq. No.
                    15053 1.R1040
 Contig ID
                    95342\overline{5}22
 5'-most EST
                    BLASTX
 Method
                    q3204108
 NCBI GI
 BLAST score
                    606
 E value
                    8.0e-63
 Match length
                    163
 % identity
                    (AJ006764) putative deoxycytidylate deaminase [Cicer
```

1616



arietinum]

Seq. No. 9033

Contig ID 15053_2.R1040

5'-most EST LIB3049-052-Q1-E1-D1

Method BLASTX
NCBI GI g3204108
BLAST score 247
E value 4.0e-21
Match length 69
% identity 64

NCBI Description (AJ006764) putative deoxycytidylate deaminase [Cicer

arietinum]

Seq. No. 9034

Contig ID 15053_3.R1040

5'-most EST LIB3050-008-Q1-E1-F5

Seq. No. 9035

Contig ID 15054_1.R1040

5'-most EST LIB3139-038-P1-N1-G6

Method BLASTX
NCBI GI g3355471
BLAST score 218
E value 3.0e-17
Match length 96
% identity 43

NCBI Description (AC004218) putative lysophospholipase [Arabidopsis

thaliana]

Seq. No. 9036

Contig ID 15058_1.R1040

5'-most EST LIB3073-004-Q1-K1-E12

Seq. No. 9037

Contig ID 15059_1.R1040

5'-most EST LIB3170-035-Q1-K1-A10

Seq. No. 9038

Contig ID 15069_1.R1040

5'-most EST LIB3049-052-Q1-E1-A7

Method BLASTX
NCBI GI g2244989
BLAST score 421
E value 3.0e-41
Match length 121
% identity 68

NCBI Description (Z97340) strong similarity to naringenin 3-dioxygenase

[Arabidopsis thaliana]

Seq. No. 9039

Contig ID 15076_1.R1040

5'-most EST LIB3049-052-Q1-E1-B4

Method BLASTX
NCBI GI g3763920
BLAST score 311
E value 1.0e-28



Match length 133 % identity 50

NCBI Description (AC004450) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9040

Contig ID 15079 1.R1040 5'-most EST kl1701210864.h1

Method BLASTX
NCBI GI g2642158
BLAST score 568
E value 4.0e-58
Match length 271
% identity 46

NCBI Description (AC003000) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9041

Contig ID 15082 1.R1040

5'-most EST LIB3049-052-Q1-E1-C12

Method BLASTX
NCBI GI g267131
BLAST score 554
E value 3.0e-56
Match length 286
% identity 27

NCBI Description NUCLEOLYSIN TIAR (TIA-1 RELATED PROTEIN)

>gi_423120_pir__A46174 RNA-binding protein TIAR - human
>gi_189310 (M96954) nucleolysin TIAR [Homo sapiens]
>gi_4507499_ref_NP_003243.1_pTIAL1_ TIA1 cytotoxic

granule-associated RNA-binding protein-like

Seq. No. 9042

Contig ID 15082 2.R1040

Seq. No. 9043

Contig ID 15083 1.R1040

5'-most EST LIB3055-011-Q1-N1-H4

Method BLASTX
NCBI GI g2496789
BLAST score 1250
E value 1.0e-138
Match length 381
% identity 64

NCBI Description 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP

REDUCTOISOMERASE) >gi_1001556_dbj_BAA10183_ (D64000)

hypothetical protein [Synechocystis sp.]

Seq. No. 9044

Contig ID 15083_2.R1040 5'-most EST seb700652689.h1

Method BLASTX
NCBI GI g2496789
BLAST score 557
E value 7.0e-57
Match length 162
% identity 65

NCBI Description 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP



REDUCTOISOMERASE) >gi_1001556_dbj_BAA10183_ (D64000)
hypothetical protein [Synechocystis sp.]

Seq. No. 9045

Contig ID 15083_3.R1040 5'-most EST uC-gmropic046g01b1

Method BLASTX
NCBI GI g2496789
BLAST score 199
E value 3.0e-15
Match length 73
% identity 55

NCBI Description 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP

REDUCTOISOMERASE) >gi_1001556_dbj_BAA10183_ (D64000)

hypothetical protein [Synechocystis sp.]

Seq. No. 9046

Contig ID 15083_4.R1040

5'-most EST LIB3139-023-P1-N1-G8

Method BLASTN
NCBI GI g2656029
BLAST score 47
E value 3.0e-17
Match length 127
% identity 84

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MQB2

Seq. No. 9047

Contig ID 15086_1.R1040 5'-most EST fC-gmse7000751793r1

Seq. No. 9048

Contig ID 15089_1.R1040

5'-most EST LIB3049-051-Q1-E1-H12

Method BLASTX
NCBI GI g1213629
BLAST score 149
E value 1.0e-09
Match length 99
% identity 41

NCBI Description (X95991) pectinesterase [Prunus persica]

Seq. No. 9049

Contig ID 15090 1.R1040

5'-most EST LIB3170-025-Q1-K1-A11

Method BLASTX
NCBI GI g1420887
BLAST score 184
E value 2.0e-13
Match length 75
% identity 44

NCBI Description (U34334) non-specific lipid transfer-like protein

[Phaseolus vulgaris]

Seq. No. 9050

Contig ID 15094 1.R1040

Seq. No.

Contig ID 5'-most EST

9058

15115 2.R1040

gsv701050017.h1



```
LIB3049-051-Q1-E1-H6
5'-most EST
                  BLASTX
Method
                  q1346478
NCBI GI
BLAST score
                  237
                   4.0e-20
E value
                   76
Match length
                   58
% identity
                  MAJOR ALLERGEN MAL D 1 (MAL D I) >gi_747852_emb_CAA58646_
NCBI Description
                   (X83672) Mal d 1 [Malus domestica]
                   9051
Seq. No.
                   15101 1.R1040
Contig ID
                   LIB3074-004-Q1-K1-G11
5'-most EST
                   BLASTX
Method
NCBI GI
                   a3702962
                   553
BLAST score
                   1.0e-56
E value
                   110
Match length
                   95
% identity
                   (AF079484) rac GTP binding protein Arac7 [Arabidopsis
NCBI Description
                   thaliana]
                   9052
Seq. No.
                   15102 1.R1040
Contig ID
                   LIB3049-052-Q1-E1-A2
5'-most EST
                   BLASTX
Method
                   q4309969
NCBI GI
                   386
BLAST score
                   4.0e-37
E value
                   128
Match length
% identity
                   58
                   (AC002983) putative phosphoglyceride transfer protein
NCBI Description
                   [Arabidopsis thaliana]
                   9053
Seq. No.
                   15104 1.R1040
Contig ID
                   LIB3049-052-Q1-E1-A4
5'-most EST
                   9054
Seq. No.
                   15104 3.R1040
Contig ID
                   trc700561613.h1
5'-most EST
                   9055
Seq. No.
                   15104 4.R1040
Contig ID
                   sat701002972.h1
 5'-most EST
 Seq. No.
                   9056
                   15107 1.R1040
 Contig ID
 5'-most EST
                   jC-gmro02910026b02a1
                   9057
 Seq. No.
                   15115 1.R1040
 Contig ID
                   trc700565473.hl
 5'-most EST
```



9059 Seq. No.

15116 1.R1040 Contig ID

LIB3049-051-Q1-E1-G11 5'-most EST

9060 Seq. No.

15117 1.R1040 Contig ID

LIB3170-032-Q1-K1-E12 5'-most EST

9061 Seq. No.

15122 1.R1040 Contig ID epx701104947.hl 5'-most EST

BLASTX Method q3695023 NCBI GI BLAST score 149 1.0e-09 E value 46 Match length 57 % identity

NCBI Description (AF055850) unknown [Arabidopsis thaliana]

9062 Seq. No.

15127 1.R1040 Contig ID

LIB3072-051-Q1-E1-G1 5'-most EST

9063 Seq. No.

15127 2.R1040 Contig ID

LIB3072-043-Q1-K1-E10 5'-most EST

9064 Seq. No.

15132 1.R1040 Contig ID pxt700945365.hl 5'-most EST

BLASTX Method g1706319 NCBI GI BLAST score 1231 E value 1.0e-135 393 Match length 58 % identity

HISTIDINE DECARBOXYLASE (HDC) (TOM92) NCBI Description

>gi_481829_pir__S39554 histidine decarboxylase (EC 4.1.1.22) - tomato >gi_416534_emb_CAA50719_ (X71900) histidine decarboxylase [Lycopersicon esculentum]

9065 Seq. No.

15132 2.R1040 Contig ID 5'-most EST leu701149945.h1

9066 Seq. No.

15135 1.R1040 Contig ID

q4396421 5'-most EST

9067 Seq. No.

15141 1.R1040 Contig ID

LIB3109-028-Q1-K1-H9 5'-most EST

BLASTN Method NCBI GI q3821780 BLAST score 33

7.0e-09 E value



Match length 37 % identity 97

NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No. 9068

Contig ID 15141_2.R1040

5'-most EST LIB3049-051-Q1-E1-D12

Method BLASTX
NCBI GI g3036796
BLAST score 480
E value 4.0e-48
Match length 106
% identity 86

NCBI Description (AL022373) putative protein [Arabidopsis thaliana]

>gi_3805858_emb_CAA21478_ (AL031986) putative protein

[Arabidopsis thaliana]

Seq. No. 9069

Contig ID 15147_1.R1040

5'-most EST LIB3049-051-Q1-E1-B9

Seq. No. 9070

Contig ID 15147_2.R1040 5'-most EST uC-gmropic031d03b1

Seq. No. 9071

Contig ID 15147 3.R1040

5'-most EST LIB3170-051-Q1-K1-H5

Seq. No. 9072

Contig ID 15148_1.R1040

5'-most EST LIB3049-051-Q1-E1-C1

Seq. No. 9073

Contig ID 15159_1.R1040

5'-most EST LIB3109-015-Q1-K1-G1

Seq. No. 9074

Contig ID 15166_1.R1040

5'-most EST uC-gmrominsoy173d03b1

Method BLASTX
NCBI GI g3036807
BLAST score 482
E value 3.0e-48
Match length 225
% identity 51

NCBI Description (AL022373) putative protein [Arabidopsis thaliana]

Seq. No. 9075

Contig ID 15167_1.R1040

5'-most EST LIB3170-035-Q1-K1-H8

Seq. No. 9076

Contig ID 15168_1.R1040 5'-most EST uC-gmropic019a03b1

Method BLASTX NCBI GI g2194132



BLAST score 939 E value 1.0e-101 Match length 468 % identity 49

NCBI Description (AC002062) No definition line found [Arabidopsis thaliana]

Seq. No. 9077

Contig ID 15169 1.R1040

5'-most EST LIB3049-051-Q1-E1-A2

Seq. No. 9078

Contig ID 15170 1.R1040

5'-most EST LIB3049-051-Q1-E1-A3

Method BLASTX
NCBI GI g3075399
BLAST score 145
E value 4.0e-09
Match length 80
% identity 44

NCBI Description (AC004484) SF16-like protein [Arabidopsis thaliana]

Seq. No. 9079

Contig ID 15175 1.R1040

5'-most EST LIB3167-019-P1-K1-C7

Method BLASTX
NCBI GI 9481384
BLAST score 158
E value 1.0e-10
Match length 72
% identity 44

NCBI Description outer envelope protein omp24, chloroplast - spinach >qi 1076265 pir S46542 chloroplast outer envelope 24 kD

protein - spinach >qi 414200 emb CAA53243 (X75563)

chloroplast outer envelope 24 kD protein (omp24) [Spinacia

oleracea]

Seq. No. 9080

Contig ID 15175 2.R1040

5'-most EST LIB30 $\overline{5}$ 0-007-Q1-E1-C4

Seq. No. 9081

Contig ID 15179_1.R1040

5'-most EST LIB3170-033-Q1-K1-C9

Seq. No. 9082

Contig ID 15185_1.R1040 5'-most EST awf700841540.h1

Method BLASTX
NCBI GI 94006893
BLAST score 533
E value 2.0e-54
Match length 126
% identity 76

NCBI Description (Z99708) aminopeptidase-like protein [Arabidopsis thaliana]

Seq. No. 9083

Contig ID 15185 2.R1040



5'-most EST kll701211190.hl
Method BLASTX

NCBI GI g4006893 BLAST score 273 E value 2.0e-24 Match length 67 % identity 66

NCBI Description (Z99708) aminopeptidase-like protein [Arabidopsis thaliana]

Seq. No. 9084

Contig ID 15187 1.R1040

5'-most EST uC-gmflminsoy030f07b1

Method BLASTX
NCBI GI g4512687
BLAST score 664
E value 1.0e-69
Match length 163
% identity 80

NCBI Description (AC006931) floral homeotic protein AGL5 [Arabidopsis

thaliana]

Seq. No. 9085

Contig ID 15194_1.R1040 5'-most EST kl1701206523.h1

Method BLASTX
NCBI GI g1871577
BLAST score 468
E value 1.0e-46
Match length 193
% identity 51

NCBI Description (Y11553) putative 21kD protein precursor [Medicago sativa]

Seq. No. 9086

Contig ID 15198 1.R1040

5'-most EST LIB3049-024-Q1-E1-A3

Method BLASTX
NCBI GI g2290528
BLAST score 1143
E value 0.0e+00
Match length 344
% identity 92

NCBI Description (U94746) ATAN11 [Arabidopsis thaliana]

Seq. No. 9087

15198 2.R1040 Contig ID 5'-most EST g4303493 Method BLASTX NCBI GI g2290532 BLAST score 829 E value 6.0e-89 Match length 183 86 % identity

NCBI Description (U94748) AN11 [Petunia x hybrida]

Seq. No. 9088

Contig ID 15203_1.R1040 5'-most EST gsv70\overline{1}051022.h1



Method BLASTX NCBI GI g3929368 BLAST score 717 1.0e-75 E value Match length 186 75 % identity

PRE-MRNA SPLICING FACTOR SF2 (SR1 PROTEIN) >gi 2443472 NCBI Description

(AF001035) ASF/SF2 homolog [Arabidopsis thaliana]

Seq. No. 9089

Contig ID 15205_1.R1040

LIB3170-035-Q1-K1-E8 5'-most EST

Method BLASTX NCBI GI q3695392 BLAST score 628 E value 5.0e-65 Match length 186 % identity 65

NCBI Description (AF096371) No definition line found [Arabidopsis thaliana]

Seq. No.

9090 15211 1.R1040 Contig ID $asn70\overline{1}134725.h2$ 5'-most EST

Method BLASTX NCBI GI q629881 BLAST score 207 E value 5.0e-16 Match length 120 49 % identity

myosin-related protein - slime mold (Physarum polycephalum) NCBI Description

>gi_511053_emb_CAA79924_ (Z21878) myosin-related protein

[Physarum polycephalum]

Seq. No. 9091

15211 2.R1040 Contig ID

5'-most EST jC-gmle01810089g09a1

9092 Seq. No.

Contig ID 15216 1.R1040 5'-most EST hyd700725833.h1

Seq. No. 9093

15216 2.R1040 Contig ID 5'-most EST sat701013479.hl

Method BLASTX NCBI GI g452593 BLAST score 336 E value 3.0e-82 Match length 266 % identity 58

NCBI Description (D21814) ORF [Lilium longiflorum]

Seq. No.

15218 1.R1040 Contig ID

5'-most EST LIB3094-030-Q1-K1-G6

9094

Seq. No. 9095



```
15218 2.R1040
Contig ID
                  LIB3051-040-Q1-K1-G5
5'-most EST
                   9096
Seq. No.
                   15218 3.R1040
Contig ID
                  uC-gmrominsoy063b02b1
5'-most EST
Seq. No.
                   9097
                   15223 1.R1040
Contig ID
                   g5342402
5'-most EST
                  BLASTX
Method
                   q4103342
NCBI GI
                   658
BLAST score
                   5.0e-69
E value
                   142
Match length
                   91
% identity
                   (AF022377) agamous-like putative transcription factor
NCBI Description
                   [Cucumis sativus]
Seq. No.
                   9098
                   15223 2.R1040
Contig ID
                   ek170\overline{0}968115.h1
5'-most EST
                   BLASTN
Method
                   g3646325
NCBI GI
                   58
BLAST score
                   4.0e-24
E value
                   145
Match length
                   86
% identity
NCBI Description Malus domestica mRNA for MADS-box protein, MADS10
                   9099
Seq. No.
                   15226 1.R1040
Contig ID
                   epx701106881.h1
5'-most EST
                   BLASTN
Method
                   q3695062
NCBI GI
BLAST score
                   124
                   3.0e-63
E value
                   258
Match length
                   87
% identity
                   Lotus japonicus rac GTPase activating protein 3 mRNA,
NCBI Description
                   partial cds
                   9100
Seq. No.
Contig ID
                   15236 1.R1040
5'-most EST
                   LIB3049-050-Q1-E1-D1
                    9101
Seq. No.
                    15236 2.R1040
 Contig ID
 5'-most EST
                   LIB3051-018-Q1-E1-H10
                   BLASTX
Method
```

Method BLASTX
NCBI GI g3702323
BLAST score 691
E value 1.0e-72
Match length 264
% identity 57

NCBI Description (AC005397) unknown protein [Arabidopsis thaliana]

```
Seq. No.
                   9102
                   15237 1.R1040
Contig ID
                   LIB31\overline{0}9-010-Q1-K1-C5
5'-most EST
                   9103
Seq. No.
                   15238 1.R1040
Contig ID
                   LIB3170-034-Q1-K1-A8
5'-most EST
Seq. No.
                   9104
                   15239 1.R1040
Contig ID
                   LIB3170-035-Q1-K1-B8
5'-most EST
                    9105
Seq. No.
                    15248 1.R1040
Contig ID
                   LIB3049-050-Q1-E1-A12
5'-most EST
                   BLASTX
Method
NCBI GI
                    g3063451
BLAST score
                    451
                    9.0e-45
E value
```

Match length 194
% identity 48
NCRI Description (ACON3981) F22013.13 [Arabidopsis the content of the content of

NCBI Description (AC003981) F22013.13 [Arabidopsis thaliana]

 Seq. No.
 9106

 Contig ID
 15251_1.R1040

 5'-most EST
 LIB3049-050-Q1-E1-A6

 Method
 BLASTX

 NCBI GI
 g4406775

 BLAST score
 363

BLAST score 363
E value 1.0e-60
Match length 358
% identity 40

NCBI Description (AC006836) unknown protein [Arabidopsis thaliana]

Seq. No. 9107 Contig ID 15253 1.R1040

5'-most EST LIB3056-001-Q1-B1-G5

Method BLASTX
NCBI GI g3132470
BLAST score 269

E value 1.0e-23
Match length 86
% identity 57

NCBI Description (AC003096) unknown protein [Arabidopsis thaliana]

Seq. No. 9108

Contig ID 15255_1.R1040 5'-most EST leu701148356.h1

Method BLASTX
NCBI GI g3269297
BLAST score 207
E value 6.0e-16
Match length 83
% identity 47

NCBI Description (AL030978) putative protein [Arabidopsis thaliana]

Seq. No. 9109



15255 2.R1040 Contig ID $uaw70\overline{0}661087.h1$ 5'-most EST BLASTX Method q3269297 NCBI GI 220 BLAST score 6.0e-18 E value 82 Match length 50 % identity NCBI Description (AL030978) putative protein [Arabidopsis thaliana] 9110 Seq. No. 15259 1.R1040 Contig ID LIB3170-033-Q1-K1-E8 5'-most EST 9111 Seq. No. 15261 1.R1040 Contig ID LIB3107-058-Q1-K1-C3 5'-most EST BLASTX Method NCBI GI g3850583 BLAST score 393 6.0e-38 E value Match length 107 % identity 71 (AC005278) Contains similarity to transcription initiation NCBI Description factor IIE, alpha subunit gb_X63468 from Homo sapiens. [Arabidopsis thaliana] 9112 Seq. No. 15265 1.R1040 Contig ID LIB3056-006-Q1-N1-H5 5'-most EST BLASTX Method g1755162 NCBI GI 715 BLAST score 2.0e-75 E value 212 Match length 67 % identity (U75192) germin-like protein [Arabidopsis thaliana] NCBI Description 9113 Seq. No. 15265 2.R1040 Contig ID LIB3170-034-Q1-K1-H7 5'-most EST BLASTX Method g1070358 NCBI GI 297 BLAST score 8.0e-27 E value 88 Match length 68 % identity (X93171) oxalate oxidase-like protein or germin-like NCBI Description protein [Hordeum vulgare] 9114 Seq. No. 15265 4.R1040 Contig ID

 Seq. No.
 9114

 Contig ID
 15265_4.R1040

 5'-most EST
 pcp700994150.h1

 Method
 BLASTX

 NCBI GI
 g2129594

 PLAST score
 223

BLAST score 223 E value 2.0e-18



Match length 70 % identity 60

NCBI Description germin type 2 - Arabidopsis thaliana

>gi_1107491_emb_CAA63023_ (X91957) germin type2

[Arabidopsis thaliana]

Seq. No. 9115

Contig ID 15266 1.R1040

5'-most EST jC-gmst02400053e08d1

9116

Seq. No.

Contig ID 15270 1.R1040

5'-most EST LIB3072-057-Q1-K1-F6

Method BLASTX
NCBI GI g3786009
BLAST score 218
E value 2.0e-17
Match length 81
% identity 57

NCBI Description (AC005499) unknown protein [Arabidopsis thaliana]

Seq. No. 9117

Contig ID 15272 1.R1040

5'-most EST LIB3107-033-Q1-K1-C1

Method BLASTN
NCBI GI g2224730
BLAST score 256
E value 1.0e-141
Match length 556
% identity 89

NCBI Description Vigna radiata mRNA for Aux22d, complete cds

Seq. No. 9118

Contig ID 15272_2.R1040 5'-most EST kl1701212274.h1

Method BLASTN
NCBI GI g2224730
BLAST score 83
E value 6.0e-39
Match length 159
% identity 88

NCBI Description Vigna radiata mRNA for Aux22d, complete cds

Seq. No. 9119

Contig ID 15274_1.R1040 5'-most EST gsv701044983.h1

Method BLASTX
NCBI GI g1703574
BLAST score 470
E value 7.0e-47
Match length 185
% identity 47

NCBI Description (U80437) C43E11.9 [Caenorhabditis elegans]

Seq. No.

9120

Contig ID 15281_1.R1040

5'-most EST LIB3050-027-Q1-E1-H2



```
BLASTX
Method
NCBI GI
                  g2598597
                  249
BLAST score
                   5.0e-21
E value
                  101
Match length
                   50
% identity
NCBI Description (Y15371) MtN5 [Medicago truncatula]
Seq. No.
                   9121
                   15281 2.R1040
Contig ID
                  LIB3050-006-Q1-E1-F5
5'-most EST
                   BLASTX
Method
                   q2598597
NCBI GI
                   315
BLAST score
                   9.0e-29
E value
                   101
Match length
                   58
% identity
NCBI Description (Y15371) MtN5 [Medicago truncatula]
                   9122
Seq. No.
                   15281_3.R1040
Contig ID
                   g5688<u>2</u>80
5'-most EST
                   BLASTX
Method
                   q2598597
NCBI GI
                   296
BLAST score
                   1.0e-26
E value
                   102
Match length
                   59
% identity
NCBI Description (Y15371) MtN5 [Medicago truncatula]
                   9123
Seq. No.
                   15281 5.R1040
Contig ID
                   LIB3107-022-Q1-K1-G6
5'-most EST
                   BLASTX
Method
                   q2598597
NCBI GI
BLAST score
                   159
                   8.0e-11
E value
Match length
                   44
% identity
                   66
NCBI Description (Y15371) MtN5 [Medicago truncatula]
                   9124
Seq. No.
Contig ID
                   15281 6.R1040
                   uC-gmropic033g05b1
5'-most EST
Seq. No.
                   9125
                   15284 1.R1040
Contig ID
5'-most EST
                   LIB3049-049-Q1-E1-F10
                   9126
Seq. No.
                   15285 1.R1040
Contig ID
5'-most EST
                   k11701208201.h1
Method
                   BLASTN
NCBI GI
                   q1420935
```

243

472

1.0e-134

BLAST score E value

Match length



% identity 88
NCBI Description Vigna unguiculata aspartic proteinase mRNA, complete cds

Seq. No. 9127

Contig ID

15291 1.R1040

5'-most EST

LIB3170-035-Q1-K1-G7

Seq. No.

9128

Contig ID

15296_1.R1040

5'-most EST

LIB3093-051-Q1-K1-E12

Method BLASTX
NCBI GI g2160183
BLAST score 431
E value 3.0e-42
Match length 134
% identity 67

% identity

NCBI Description

(AC000132) Identical to A. thaliana U2 SnRNP-specific A' protein (gb_X69137). ESTs gb_ATTS0705, gb_ATTS0339 come

from this gene. [Arabidopsis thaliana]

Seq. No. 9129

Contig ID 15300_1.R1040 5'-most EST taw700654794.h1

Method BLASTX
NCBI GI g2459442
BLAST score 360
E value 1.0e-33
Match length 223
% identity 43

NCBI Description (AC002332) putative DNA-binding protein PD1 [Arabidopsis

thaliana]

Seq. No. 9130

Contig ID 15301_1.R1040

5'-most EST uC-gmronoir037e10b1

Method BLASTX
NCBI GI g2281633
BLAST score 473
E value 6.0e-47
Match length 184
% identity 57

NCBI Description (AF003097) AP2 domain containing protein RAP2.4

[Arabidopsis thaliana]

Seq. No. 9131

Contig ID 15301_2.R1040 5'-most EST g5058437 Method BLASTX

NCBI GI g2281633 BLAST score 428 E value 2.0e-50 Match length 180

% identity 64

NCBI Description (AF003097) AP2 domain containing protein RAP2.4

[Arabidopsis thaliana]

Seq. No. 9132



Contig ID 15301_3.R1040 5'-most EST gsv701052436.h1

Seq. No. 9133

Contig ID 15317_1.R1040

5'-most EST LIB $30\overline{4}9-049-Q1-E1-A6$

Seq. No. 9134

Contig ID 15318 1.R1040

5'-most EST LIB3049-049-Q1-E1-A7

Method BLASTX
NCBI GI g1730109
BLAST score 382
E value 5.0e-37
Match length 99
% identity 81

NCBI Description LEUCOANTHOCYANIDIN DIOXYGENASE (LDOX) (LEUCOANTHOCYANIDIN

HYDROXYLASE) >gi_499022_emb_CAA53580_ (X75966) leucoanthocyanidin dioxygenase [Vitis vinifera]

Seq. No. 9135

Contig ID 15325 1.R1040

5'-most EST LIB3170-036-Q1-K1-E7

Seq. No. 9136

Contig ID 15328 1.R1040

5'-most EST LIB3093-036-Q1-K1-D8

Seq. No. 9137

Contig ID 15329 1.R1040

5'-most EST LIB3092-014-Q1-K1-C10

Method BLASTX
NCBI GI g3434971
BLAST score 307
E value 2.0e-27
Match length 187

% identity 45

NCBI Description (AB008105) ethylene responsive element binding factor 3

íz-

[Arabidopsis thaliana]

Seq. No. 9138

Contig ID 15329_2.R1040

5'-most EST LIB3049-049-Q1-E1-B9

Method BLASTX
NCBI GI g3434971
BLAST score 237
E value 1.0e-19
Match length 67
% identity 69

NCBI Description (AB008105) ethylene responsive element binding factor 3

[Arabidopsis thaliana]

Seq. No. 9139

Contig ID 15331_1.R1040 5'-most EST ssr700559243.h1

Seq. No. 9140



Contig ID 15331_2.R1040 5'-most EST vzy700752955.h1

Seq. No. 9141

Contig ID 15335_1.R1040 5'-most EST awf700838743.h1

Method BLASTX
NCBI GI g4490339
BLAST score 190
E value 1.0e-13
Match length 325
% identity 27

NCBI Description (AL035656) putative protein [Arabidopsis thaliana]

Seq. No. 9142

Contig ID 15335_2.R1040 5'-most EST crh700851218.h1

Seq. No. 9143

Contig ID 15335_3.R1040 5'-most EST 6HC-01-Q1-E1-H2

Seq. No. 9144

Contig ID 15335_4.R1040 5'-most EST crh700850801.h1

Method BLASTX
NCBI GI 94490339
BLAST score 148
E value 2.0e-09
Match length 80
% identity 41

NCBI Description (AL035656) putative protein [Arabidopsis thaliana]

Seq. No. 9145

Contig ID 15335_5.R1040 5'-most EST txt700732387.h1

Seq. No. 9146

Contig ID 15337_1.R1040

5'-most EST LIB3093-042-Q1-K1-C6

Method BLASTX
NCBI GI g1922242
BLAST score 175
E value 2.0e-12
Match length 70
% identity 53

NCBI Description (Y10084) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9147

Contig ID 15337_2.R1040 5'-most EST zsg701123968.h1

Method BLASTX
NCBI GI g1922242
BLAST score 210
E value 1.0e-16
Match length 70
% identity 56



NCBI Description (Y10084) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9148

Contig ID 15337_3.R1040

5'-most EST LIB3049-048-Q1-E1-H11

Seq. No. 9149

Contig ID 15339_1.R1040

5'-most EST jC-gmle01810046c08a1

Method BLASTX
NCBI GI g1730502
BLAST score 255
E value 9.0e-22
Match length 160
% identity 38

NCBI Description TRANSMEMBRANE PROTEIN PFT27 >gi_110903_pir_A31351 probable

transmembrane protein FT27 - mouse >gi_535682 (M23568)

transmembrane protein [Mus musculus]

Seq. No. 9150

Contig ID 15342_1.R1040

5'-most EST LIB3109-036-Q1-K1-F2

Method BLASTN
NCBI GI g3335331
BLAST score 41
E value 2.0e-13
Match length 189

Match length 189 % identity 80

NCBI Description Arabidopsis thaliana chromosome 1 BAC T8F5 sequence,

complete sequence [Arabidopsis thaliana]

Seq. No. 9151

Contig ID 15345_1.R1040

5'-most EST jC-gmro02800040f08a1

Method BLASTX
NCBI GI g1297187
BLAST score 373
E value 3.0e-35
Match length 170
% identity 46

NCBI Description (U53501) similar to protein encoded by GenBank Accession

Number U41815, nucleoporin 98 [Arabidopsis thaliana]

Seq. No. 9152

Contig ID 15346_1.R1040

5'-most EST LIB3170-027-Q2-K1-G11

Method BLASTX
NCBI GI g3914667
BLAST score 451
E value 1.0e-44
Match length 136
% identity 72

NCBI Description 50S RIBOSOMAL PROTEIN L28, CHLOROPLAST PRECURSOR

>gi_2459427 (AC002332) putative chloroplast 50S ribosomal

protein L28 [Arabidopsis thaliana]

Seq. No. 9153



Contig ID 15347_1.R1040 . 5'-most EST jC-gmro02910035b06d1

Seq. No. 9154

Contig ID 15348_1.R1040

5'-most EST LIB3170-033-Q1-K1-B7

Seq. No. 9155

Contig ID 15349 1.R1040

5'-most EST LIB3072-061-Q1-K1-B1

Seq. No. 9156

Contig ID 15353 1.R1040

5'-most EST LIB3049-048-Q1-E1-G9

Method BLASTX
NCBI GI g2708484
BLAST score 559
E value 1.0e-57
Match length 117
% identity 85

NCBI Description (U79557) IAA24 [Arabidopsis thaliana]

Seq. No. 9157

Contig ID 15358_1.R1040

5'-most EST LIB3049-048-Q1-E1-F7

Seq. No. 9158

Contig ID 15360 1.R1040

5'-most EST LIB3170-034-Q1-K1-D6

Method BLASTX
NCBI GI g2827555
BLAST score 421
E value 2.0e-41
Match length 139
% identity 44

NCBI Description (AL021635) Translation factor EF-1 alpha - like protein

[Arabidopsis thaliana]

Seq. No. 9159

Contig ID 15361_1.R1040 5'-most EST ncj700975258.h1

Method BLASTX
NCBI GI g3641837
BLAST score 809
E value 1.0e-86
Match length 243
% identity 70

NCBI Description (AL023094) Nonclathrin coat protein gamma - like protein

[Arabidopsis thaliana]

Seq. No. 9160

Contig ID 15363_1.R1040

5'-most EST LIB3049-048-Q1-E1-G12

Seq. No. 9161

Contig ID 15364_1.R1040

5'-most EST LIB3049-048-Q1-E1-G2



```
9162
Seq. No.
                   15367 1.R1040
Contig ID
                   uC-gmropic075g06b1
5'-most EST
                   BLASTX
Method
                   q3063457
NCBI GI
BLAST score
                   298
                   5.0e-27
E value
Match length
                   113
                   55
% identity
NCBI Description (AC003981) F22013.19 [Arabidopsis thaliana]
Seq. No.
                   9163
                   15367 2.R1040
Contig ID
5'-most EST
                   ish701064577.hl
                   BLASTX
Method
                   q3063457
NCBI GI
                   165
BLAST score
E value
                   2.0e-11
Match length
                   99
                   39
% identity
                   (AC003981) F22013.19 [Arabidopsis thaliana]
NCBI Description
                   9164
Seq. No.
                   15368 1.R1040
Contig ID
                   fua701041455.hl
5'-most EST
                   BLASTX
Method
                   q3551954
NCBI GI
BLAST score
                   607
                   9.0e-63
E value
Match length
                   245
                   47
% identity
                   (AF082030) senescence-associated protein 5 [Hemerocallis
NCBI Description
                   hybrid cultivar]
Seq. No.
                   9165
                   15374 1.R1040
Contig ID
                   LIB3049-048-Q1-E1-E1
5'-most EST
                   9166
Seq. No.
                   15375 1.R1040
Contig ID
                   LIB3049-048-Q1-E1-E10
5'-most EST
                   9167
Seq. No.
                   15377 1.R1040
Contig ID
                   LIB3053-011-Q1-N1-A4
 5'-most EST
                    9168
 Seq. No.
                   15380 1.R1040
Contig ID
                   LIB3049-011-Q1-E1-B5
 5'-most EST
                   BLASTX
Method
                   g3860274
 NCBI GI
                    338
 BLAST score
```

(AC005824) unknown protein [Arabidopsis thaliana]

4.0e-31

160

42

E value Match length

% identity

NCBI Description



>gi_4314397_gb_AAD15607_ (AC006232) putative zinc finger
protein [Arabidopsis thaliana]

Seq. No. 9169

Contig ID 15380_2.R1040 5'-most EST wrg700787816.h2

Seq. No. 9170

Contig ID 15380 8.R1040

5'-most EST LIB3049-030-Q1-E1-A4

Seq. No. 9171

Contig ID 15388_1.R1040 5'-most EST jex700909190.h1

Method BLASTX
NCBI GI g2583108
BLAST score 177
E value 1.0e-12
Match length 52
% identity 63

NCBI Description (AC002387) putative surface protein [Arabidopsis thaliana]

Seq. No. 9172

Contig ID 15395 1.R1040

5'-most EST LIB3170-036-Q1-K1-H6

Method BLASTX
NCBI GI g4544407
BLAST score 356
E value 1.0e-33
Match length 96
% identity 61

NCBI Description (AC006955) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9173

Contig ID 15397_1.R1040

5'-most EST LIB3049-048-Q1-E1-D10

Method BLASTX
NCBI GI g3914467
BLAST score 579
E value 8.0e-60
Match length 145
% identity 79

NCBI Description 26S PROTEASOME REGULATORY SUBUNIT S3 (NUCLEAR ANTIGEN 21D7)

>gi 1864003 dbj BAA19252 (AB001422) 21D7 [Nicotiana

tabacum]

Seq. No. 9174

Contig ID 15399 1.R1040

5'-most EST LIB3170-034-Q1-K1-B6

Seq. No. 9175

Contig ID 15401_1.R1040

5'-most EST LIB3049-048-Q1-E1-B4

Seq. No. 9176

Contig ID 15402_1.R1040 5'-most EST uC-gmronoir051g02b1



Method BLASTX
NCBI GI g4220474
BLAST score 1493
E value 1.0e-166
Match length 487
% identity 66

NCBI Description (AC006069) putative myosin heavy chain [Arabidopsis

thaliana]

Seq. No. 9177

Contig ID 15402 2.R1040

5'-most EST jC-gmro02800023b09d1

Seq. No. 9178

Contig ID 15404 1.R1040

5'-most EST LIB3170-036-Q1-K2-C6

Seq. No. 9179

Contig ID 15408 1.R1040

5'-most EST LIB3139-072-P1-N1-D6

Seq. No. 9180

Contig ID 15414 1.R1040

5'-most EST LIB3170-036-Q1-K2-F6

Method BLASTX
NCBI GI g2315451
BLAST score 291
E value 7.0e-26
Match length 224
% identity 31

NCBI Description (AF016448) No definition line found [Caenorhabditis

elegans]

Seq. No. 9181

Contig ID 15417_1.R1040 5'-most EST vwf700675916.h1

Method BLASTN
NCBI GI g12387
BLAST score 34
E value 2.0e-09
Match length 69
% identity 94

NCBI Description Vicia faba chloroplast tRNA-Leu, tRNA-Phe, tRNA-His, ORFX,

NADH-deydrogenase genes & partial sequence ORFx & psbA

genes

Seq. No. 9182

Contig ID 15420 1.R1040

5'-most EST LIB $30\overline{4}$ 9-047-Q1-E1-G12

Seq. No. 9183

Contig ID 15425 1.R1040

5'-most EST LIB3170-035-Q1-K1-H5

Seq. No. 9184

Contig ID 15426 1.R1040 5'-most EST vwf700677929.h1

Method

NCBI GI

BLASTX g3236240



```
BLASTX
Method
NCBI GI
                   g3236240
                   947
BLAST score
                   1.0e-102
E value
                   428
Match length
                   75
% identity
NCBI Description (AC004684) unknown protein [Arabidopsis thaliana]
                 9185
Seq. No.
                   15426 2.R1040
Contig ID
                   uC-gmropic041f01b1
5'-most EST
                   BLASTX
Method
                   q4099833
NCBI GI
                   227
BLAST score
                   3.0e-18
E value
Match length
                   70
                   61
% identity
                   (U90265) bifunctional nuclease [Zinnia elegans]
NCBI Description
                   9186
Seq. No.
                   15426 3.R1040
Contig ID
                   leu701155823.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3236240
                   1071
BLAST score
                   1.0e-117
E value
                   292
Match length
% identity
                   72
                   (AC004684) unknown protein [Arabidopsis thaliana]
NCBI Description
                   9187
Seq. No.
                   15426 4.R1040
Contig ID
                   LIB3051-050-Q1-K1-B10
 5'-most EST
                   BLASTX
Method
NCBI GI
                   g3236240
BLAST score
                   373
                   1.0e-35
 E value
                   77
Match length
                   91
 % identity
                    (AC004684) unknown protein [Arabidopsis thaliana]
 NCBI Description
                    9188
 Seq. No.
                    15426 5.R1040
 Contig ID
                    zlv700807524.hl
 5'-most EST
                   BLASTX
 Method
                    g3236240
 NCBI GI
                    604
 BLAST score
                    1.0e-74
 E value
                    175
 Match length
                    80
 % identity
                   (AC004684) unknown protein [Arabidopsis thaliana]
 NCBI Description
                    9189
 Seq. No.
                    15426 9.R1040
 Contig ID
                    LIB3051-019-Q1-E1-C6
 5'-most EST
```



BLAST score 213 E value 2.0e-22 Match length 77

% identity 71

NCBI Description (AC004684) unknown protein [Arabidopsis thaliana]

Seq. No.

9190

Contig ID

15436 1.R1040

5'-most EST

LIB3049-044-Q1-E1-E1

Seq. No.

9191

Contig ID

15436 2.R1040

5'-most EST

LIB3109-022-Q1-K1-E12

Seq. No.

9192

Contig ID

15439 1.R1040

5'-most EST

uC-gmflminsoy011d09b1

Seq. No.

9193

Contig ID

15441 1.R1040

5'-most EST

LIB3073-006-Q1-K1-F10

Seq. No.

9194

Contig ID

15443 1.R1040

5'-most EST

LIB3049-047-Q1-E1-F1

Seq. No.

Contig ID

15445 1.R1040

5'-most EST

LIB3051-003-Q1-E1-F12

Method NCBI GI BLASTX g2129541

BLAST score E value

1586 1.0e-177

Match length

474

9195

% identity

67

NCBI Description

ATPK19 protein - Arabidopsis thaliana >gi 914079 bbs 160872

ATPK19=ribosomal-protein S6 kinase homolog [Arabidopsis

thaliana, Peptide, 471 aa] >gi 867995 dbj BAA07661

(D42061) ribosomal-protein S6 kinase homolog [Arabidopsis

thaliana]

Seq. No.

9196

Contig ID

15445 3.R1040

5'-most EST

jC-qmle01810020b07d1

Method

BLASTX

NCBI GI

g2129541 219

BLAST score E value

9.0e-18

Match length

63 65

% identity NCBI Description

ATPK19 protein - Arabidopsis thaliana >gi_914079_bbs_160872

ATPK19=ribosomal-protein S6 kinase homolog [Arabidopsis

thaliana, Peptide, 471 aa] >qi 867995 dbj BAA07661

(D42061) ribosomal-protein S6 kinase homolog [Arabidopsis

thaliana]

Seq. No.



Contig ID 15449 1.R1040

5'-most EST LIB30 $\overline{4}$ 9-047-Q1-E1-F7

Method BLASTX
NCBI GI g1731146
BLAST score 406
E value 3.0e-39
Match length 186
% identity 45

NCBI Description HYPOTHETICAL 42.1 KD PROTEIN ZK1307.9 IN CHROMOSOME III

>gi_3881615_emb_CAA87435_ (Z47358) similar to YJU2 protein

[Caenorhabditis elegans]

Seq. No. 9198

Contig ID 15449_2.R1040 5'-most EST sat701011550.h1

Method BLASTX
NCBI GI g1731146
BLAST score 209
E value 8.0e-17
Match length 54
% identity 69

NCBI Description HYPOTHETICAL 42.1 KD PROTEIN ZK1307.9 IN CHROMOSOME III

>gi_3881615_emb_CAA87435_ (Z47358) similar to YJU2 protein

[Caenorhabditis elegans]

Seq. No. 9199

Contig ID 15451_1.R1040 5'-most EST leu701156291.h1

Seq. No. 9200

Contig ID 15451 2.R1040

5'-most EST LIB3049-047-Q1-E1-G1

Method BLASTX
NCBI GI g4406775
BLAST score 429
E value 1.0e-41
Match length 379
% identity 33

NCBI Description (AC006836) unknown protein [Arabidopsis thaliana]

Seq. No. 9201

Contig ID 15454 1.R1040

5'-most EST LIB3139-004-P1-N1-E1

Method BLASTX
NCBI GI g4127781
BLAST score 352
E value 4.0e-33
Match length 130
% identity 22

NCBI Description (AJ012588) Notchless protein [Drosophila melanogaster]

Seq. No. 9202

Contig ID 15454 2.R1040

5'-most EST uC-gmflminsoy007g09b1

Seq. No. 9203

Contig ID 15455 1.R1040



```
5'-most EST LIB3170-033-Q1-K1-F5
```

Seq. No. 9204

Contig ID 15456 1.R1040

5'-most EST LIB3087-002-Q1-K1-A7

Method BLASTX
NCBI GI g3600051
BLAST score 185
E value 2.0e-13
Match length 44
% identity 77

NCBI Description (AF080120) contains similarity to the single-strand binding

proteins family (Pfam: SSB.hmm, score: 24.02) [Arabidopsis

thaliana]

Seq. No. 9205

Contig ID 15463 1.R1040

5'-most EST LIB3049-047-Q1-E1-D11

Method BLASTX
NCBI GI g4138581
BLAST score 275
E value 2.0e-24
Match length 131
% identity 21

NCBI Description (X98474) mitochondrial energy transfer protein [Solanum

tuberosum]

Seq. No. 9206

Contig ID 15466 1.R1040

5'-most EST LIB3092-011-Q1-K1-C5

Method BLASTN
NCBI GI g3985934
BLAST score 53
E value 8.0e-21
Match length 279
% identity 83

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MJE7, complete sequence [Arabidopsis thaliana]

Seq. No. 9207

Contig ID 15467 1.R1040

5'-most EST LIB3049-047-Q1-E1-D5

Seq. No. 9208

Contig ID 15476_1.R1040

5'-most EST LIB3170-033-Q1-K1-E5

Method BLASTX
NCBI GI g1168251
BLAST score 416
E value 9.0e-41
Match length 95
% identity 79

NCBI Description PROBABLE CYSTEINE PROTEINASE A494 PRECURSOR

>gi_1076384_pir__S46535 probable cysteine proteinase (EC
3.4.22.-) (clone A1494) - Arabidopsis thaliana (fragment)
>gi_516865_emb_CAA52403_ (X74359) putative thiol protease

[Arabidopsis thaliana]



Seq. No. 9209

Contig ID 15477 1.R1040

5'-most EST LIB3170-036-Q1-K2-F5

Method BLASTX
NCBI GI g4006904
BLAST score 278
E value 1.0e-24
Match length 103
% identity 54

NCBI Description (Z99708) putative MADS-box protein [Arabidopsis thaliana]

Seq. No. 9210

Contig ID 15483_1.R1040

5'-most EST LIB3049-047-Q1-E1-B7

Seq. No. 9211

Contig ID 15483_2.R1040 5'-most EST zzp700835537.h1

Seq. No. 9212

Contig ID 15484 1.R1040

5'-most EST LIB3051-039-Q1-K1-B12

Seq. No. 9213

Contig ID 15484 2.R1040 5'-most EST bth700847209.h1

Seq. No. 9214

Contig ID 15485 1.R1040

5'-most EST LIB3049-047-Q1-E1-B9

Seq. No. 9215

Contig ID 15485_2.R1040 5'-most EST bth700845065.h1

Seq. No. 9216

Contig ID 15486_1.R1040

5'-most EST LIB3049-047-Q1-E1-C1

Method BLASTX
NCBI GI g3859591
BLAST score 202
E value 2.0e-15
Match length 104
% identity 40

NCBI Description (AF104919) No definition line found [Arabidopsis thaliana]

Seq. No. 9217

Contig ID 15489_1.R1040 5'-most EST zpv700763607.h1

Seq. No. 9218

Contig ID 15492 1.R1040 5'-most EST g5126671 Method BLASTX NCBI GI g4538978

BLAST score 616



E value 8.0e-64 Match length 163 72 % identity

NCBI Description (AL049487) hypothetical protein [Arabidopsis thaliana]

9219 Seq. No.

15499 1.R1040 Contig ID hyd700725239.h1 5'-most EST

Method BLASTX NCBI GI g2911047 BLAST score 400 E value 8.0e-39 Match length 112 % identity 73

(AL021961) putative protein [Arabidopsis thaliana] NCBI Description

Seq. No.

9220

Contig ID

15500 1.R1040

5'-most EST LIB3109-022-Q1-K1-C2

Seq. No.

9221

Contig ID 15500 2.R1040

LIB3051-104-Q1-K1-C6 5'-most EST

Seq. No.

9222 Contig ID 15500 3.R1040 5'-most EST fC-gmf1700907636a1

Seq. No.

9223

15507 1.R1040 Contig ID

5'-most EST LIB3049-046-Q1-E1-F5

Method BLASTX NCBI GI g4335725 BLAST score 236 8.0e-20 E value Match length 85

% identity 56

NCBI Description (AC006248) hypothetical protein [Arabidopsis thaliana]

9224 Seq. No.

15513 1.R1040 Contig ID

5'-most EST LIB3138-090-P1-N1-A6

Method BLASTX NCBI GI g2827699 BLAST score 447 6.0e-44 E value Match length 278 44 % identity

NCBI Description (AL021684) predicted protein [Arabidopsis thaliana]

Seq. No.

9225

Contig ID 15515 1.R1040

5'-most EST LIB3170-033-Q1-K1-A5

Seq. No.

9226

Contig ID 15516 1.R1040

5'-most EST

LIB3049-046-Q1-E1-G4



```
9227
Seq. No.
                  15518 1.R1040
Contig ID
                  LIB3049-046-Q1-E1-G7
5'-most EST
                  BLASTN
Method
                  q3421089
NCBI GI
                  270
BLAST score
                  1.0e-150
E value
                  654
Match length
% identity
NCBI Description Arabidopsis thaliana 20S proteasome subunit PAE2 (PAE2)
                  mRNA, complete cds
                  9228
Seq. No.
                  15518 2.R1040
Contig ID
                  LIB3170-052-Q1-J1-F7
5'-most EST
                   9229
Seq. No.
                   15518 3.R1040
Contig ID
                   LIB3051-029-Q1-K1-B3
5'-most EST
                   BLASTN
Method
                   g3421089
NCBI GI
                   125
BLAST score
                   8.0e-64
E value
                   333
Match length
% identity
                   Arabidopsis thaliana 20S proteasome subunit PAE2 (PAE2)
NCBI Description
                   mRNA, complete cds
                   9230
Seq. No.
                   15518 4.R1040
Contig ID
                   fua701039082.h1
5'-most EST
                   BLASTX
Method
                   g3421090
NCBI GI
                   224
BLAST score
                   3.0e-36
E value
Match length
                   101
                   79
% identity
                   (AF043525) 20S proteasome subunit PAE2 [Arabidopsis
NCBI Description
                   thaliana]
                   9231
 Seq. No.
                   15523 1.R1040
 Contig ID
                   uC-qmflminsoy018c06b1
 5'-most EST
                   BLASTX
Method
 NCBI GI
                   g2129559
 BLAST score
                   754
                   3.0e-80
 E value
                   181
 Match length
                   80
 % identity
                   cellulase homolog OR16pep - Arabidopsis thaliana
 NCBI Description
                   >gi_1022807 (U37702) cellulase [Arabidopsis thaliana]
                   >gi_3493633 (AF074092) cellulase [Arabidopsis thaliana]
                   >gi_3598956 (AF074375) cellulase [Arabidopsis thaliana]
                   >gi 3978258 (AF073875) endo-1,4-beta-D-glucanase KORRIGAN
```

[Arabidopsis thaliana]



Seq. No. 9232

Contig ID 15523_2.R1040 5'-most EST pmv700890951.h1

Method BLASTX
NCBI GI g2129559
BLAST score 162
E value 1.0e-20
Match length 70
% identity 79

% identity 79 NCBI Description cellulase homolog OR16pep - Arabidopsis thaliana

>gi_1022807 (U37702) cellulase [Arabidopsis thaliana]
>gi_3493633 (AF074092) cellulase [Arabidopsis thaliana]
>gi_3598956 (AF074375) cellulase [Arabidopsis thaliana]
>gi_3978258 (AF073875) endo-1,4-beta-D-glucanase KORRIGAN

[Arabidopsis thaliana]

Seq. No. 9233

Contig ID 15525 1.R1040

5'-most EST uC-gmflminsoy063d10b1

Seq. No. 9234

Contig ID 15526 1.R1040

5'-most EST jC-qmf102220143d02d1

Method BLASTX
NCBI GI g421948
BLAST score 666
E value 9.0e-70
Match length 191
% identity 64

NCBI Description UDP rhamnose--anthocyanidin-3-glucoside

rhamnosyltransferase - garden petunia

Seq. No. 9235

Contig ID 15527_1.R1040 5'-most EST pmv700890051.h1

Seq. No. 9236

Contig ID 15528 1.R1040

5'-most EST LIB3049-046-Q1-E1-E12

Method BLASTX
NCBI GI g3600033
BLAST score 570
E value 7.0e-59
Match length 149
% identity 71

NCBI Description (AF080119) contains similarity to the N terminal domain of

the E1 protein (Pfam: E1_N.hmm, score: 12.36) [Arabidopsis

thaliana]

Seq. No. 9237

Contig ID 15529 1.R1040

5'-most EST LIB3170-035-Q1-K1-D4

Seq. No. 9238

Contig ID 15535_1.R1040 5'-most EST zhf700952064.h1

Method BLASTX



NCBI GI g1946367 BLAST score 215 E value 6.0e-17 Match length 65 % identity 62

NCBI Description (U93215) unknown protein [Arabidopsis thaliana]

Seq. No. 9239

Contig ID 15535 2.R1040

5'-most EST uC-gmflminsoy040a03b1

Method BLASTX
NCBI GI g1946367
BLAST score 611
E value 2.0e-63
Match length 164
% identity 70

NCBI Description (U93215) unknown protein [Arabidopsis thaliana]

Seq. No. 9240

Contig ID 15537 1.R1040

5'-most EST jC-gmro02910061h09a1

Method BLASTX
NCBI GI g3724328
BLAST score 144
E value 1.0e-08
Match length 160
% identity 27

NCBI Description (AB013095) heme-binding protein [Mus musculus]

Seq. No. 9241

Contig ID 15541 1.R1040

5'-most EST LIB3049-046-Q1-E1-C3

Seq. No. 9242

Contig ID 15541 2.R1040

5'-most EST LIB3072-061-Q1-K1-A2

Seq. No. 9243

Contig ID 15542 1.R1040 5'-most EST pcp700993990.h1

Seq. No. 9244

Contig ID 15546 1.R1040

5'-most EST LIB3170-028-Q1-K1-B12

Seq. No. 9245

Contig ID 15546 2.R1040 5'-most EST leu701147556.h1

Seq. No. 9246

Contig ID 15550 1.R1040

5'-most EST LIB3170-034-Q1-K1-B4

Seq. No. 9247

Contig ID 15550 2.R1040

5'-most EST LIB3138-097-Q1-N1-F10



Seq. No. 9248

Contig ID 15551 1.R1040

5'-most EST LIB3049-046-Q1-E1-B9

Method BLASTX
NCBI GI g1706311
BLAST score 376
E value 5.0e-36
Match length 189
% identity 43

NCBI Description PROBABLE ATP-DEPENDENT RNA HELICASE DBP1 (HELICASE CA1)

>gi_2132404_pir__S62003 probable ATP-dependent RNA helicase

DBPI - yeast (Saccharomyces cerevisiae) >gi_1163095 (U43503) Dbp1p,Lph8p [Saccharomyces cerevisiae]

Seq. No. 9249

Contig ID 15557 1.R1040 5'-most EST zsg701119272.h1

Seq. No. 9250

Contig ID 15560 1.R1040

5'-most EST LIB3170-033-Q1-K1-D4

Method BLASTX
NCBI GI g1076715
BLAST score 239
E value 9.0e-20
Match length 82
% identity 50

NCBI Description abscisic acid-induced protein HVA22 - barley >gi_404589

(L19119) A22 [Hordeum vulgare]

Seq. No. 9251

Contig ID 15561_1.R1040 5'-most EST fua701043583.h1

Method BLASTX
NCBI GI g2252631
BLAST score 263
E value 7.0e-23
Match length 83
% identity 63

NCBI Description (U95973) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9252

Contig ID 15561_2.R1040 5'-most EST trc700561886.h1

Seq. No. 9253

Contig ID 15565_1.R1040

5'-most EST LIB3049-046-Q1-E1-B3

Method BLASTX
NCBI GI g2288985
BLAST score 216
E value 3.0e-17
Match length 94
% identity 18

NCBI Description (AC002335) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9254



```
15566 1.R1040
Contig ID
                  LIB3049-046-Q1-E1-B4
5'-most EST
                  9255
Seq. No.
                  15572 1.R1040
Contig ID
                  hrw701060691.hl
5'-most EST
                  9256
Seq. No.
                  15579 1.R1040
Contig ID
                  LIB3170-033-Q1-K1-A4
5'-most EST
                   9257
Seq. No.
                   15583 1.R1040
Contig ID
                  LIB3170-036-Q1-K1-B4
5'-most EST
                   9258
Seq. No.
                   15587 1.R1040
Contig ID
                   LIB3049-046-Q1-E1-A2
5'-most EST
                   BLASTN
Method
                   g2959438
NCBI GI
                   79
BLAST score
                   2.0e-36
E value
                   229
Match length
% identity
                   Glycine max gene for phosphoenolpyruvate carboxylase, 5'
NCBI Description
                   untranslated region, partial cds
                   9259
Seq. No.
                   15594 1.R1040
Contig ID
                   uC-gmrominsoy298f01b1
5'-most EST
Method
                   BLASTX
                   g1246823
NCBI GI
BLAST score
                   1426
                   1.0e-158
E value
Match length
                   349
 % identity
                   80
                   (X89865) unknown [Phoenix dactylifera]
NCBI Description
                   9260
 Seq. No.
                   15594 2.R1040
 Contig ID
                   uC-gmflminsoy084e03b1
 5'-most EST
                   BLASTX
 Method
                   g1246823
 NCBI GI
                   170
 BLAST score
                   4.0e-12
 E value
                   49
 Match length
                    69
 % identity
                   (X89865) unknown [Phoenix dactylifera]
 NCBI Description
```

Seq. No. 9261

Contig ID 15595_1.R1040

5'-most EST LIB3170-034-Q1-K1-G3

Seq. No. 9262

Contig ID 15597 1.R1040

5'-most EST LIB3170-036-Q1-K1-A4

NCBI Description



```
9263
Seq. No.
                  15598 1.R1040
Contig ID
                  uC-gmrominsoy060e11b1
5'-most EST
Method
                  BLASTX
                  g4469026
NCBI GI
BLAST score
                  294
                   2.0e-26
E value
Match length
                   106
% identity
                   60
                   (AL035602) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   15598 2.R1040
Contig ID
5'-most EST
                   pcp700991182.h1
Method
                   BLASTX
NCBI GI
                   g4469026
BLAST score
                   185
                   9.0e-14
E value
                   70
Match length
% identity
                   (AL035602) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   9265
Seq. No.
                   15605 1.R1040
Contig ID
                   smc700747914.h1
5'-most EST
                   9266
Seq. No.
                   15605 2.R1040
Contig ID
                   k11701206351.h1
5'-most EST
                   9267
Seq. No.
                   15610 1.R1040
Contig ID
                   LIB3170-034-Q1-K1-F8
5'-most EST
Seq. No.
                   9268
                   15612 1.R1040
Contig ID
                   crh700852730.hl
5'-most EST
                   BLASTX
Method
                   g4263707
NCBI GI
BLAST score
                   819
                    1.0e-110
E value
                    307
Match length
 % identity
                    (AC006223) putative 70kD heat shock protein [Arabidopsis
NCBI Description
                    thaliana]
                    9269
 Seq. No.
                    15613 1.R1040
 Contig ID
                    uC-gmronoir0001c11a1
 5'-most EST
                    BLASTX
 Method
                    g4337044
 NCBI GI
                    319
 BLAST score
                    2.0e-29
 E value
                    90
 Match length
 % identity
                    69
                    (AF124161) molybdopterin synthase sulphurylase [Nicotiana
```

plumbaginifolia]



9270

Seq. No.

```
15613 2.R1040
Contig ID
                  LIB3049-045-Q1-E1-E4
5'-most EST
                  BLASTX
Method
                  g4337044
NCBI GI
                  297
BLAST score
E value
                  5.0e-27
                  86
Match length
% identity
                   65
                   (AF124161) molybdopterin synthase sulphurylase [Nicotiana
NCBI Description
                  plumbaginifolia]
                   9271
Seq. No.
                   15614 1.R1040
Contig ID
                  LIB3170-036-Q1-J1-F11
5'-most EST
Seq. No.
                   9272
                   15615 1.R1040
Contig ID
                   LIB3109-014-Q1-K1-A8
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2982243
                   475
BLAST score
                   3.0e-47
E value
Match length
                   158
% identity
                   58
                   (AF051204) hypothetical protein [Picea mariana]
NCBI Description
                   9273
Seq. No.
                   15615 2.R1040
Contig ID
                   LIB3107-029-Q1-K1-E11
5'-most EST
                   BLASTX
Method
                   g2982243
NCBI GI
BLAST score
                   187
E value
                   7.0e-14
                   69
Match length
                   54
% identity
                   (AF051204) hypothetical protein [Picea mariana]
NCBI Description
                   9274
Seq. No.
                   15615 4.R1040
Contig ID
                   LIB3109-055-Q1-K1-F5
5'-most EST
                   BLASTX
Method
                   g2982243
NCBI GI
BLAST score
                   196
                   4.0e-15
E value
                   93
Match length
                   45
% identity
                   (AF051204) hypothetical protein [Picea mariana]
NCBI Description
                   9275
Seq. No.
                   15617 1.R1040
Contig ID
                   LIB3094-011-Q1-K1-E3
5'-most EST
                   BLASTN
Method
NCBI GI
                   g21045
                   214
BLAST score
                   1.0e-116
E value
```



```
Match length
                   691
% identity
NCBI Description
                  P.vulgaris mRNA for prolin-rich protein (PvPRP1)
Seq. No.
                  9276
                  15617 2.R1040
Contig ID
5'-most EST
                  asn701135096.h1
Method
                  BLASTN
NCBI GI
                  g21045
BLAST score
                  35
E value
                  3.0e-10
Match length
                  93
% identity
                  90
NCBI Description P.vulgaris mRNA for prolin-rich protein (PvPRP1)
                  9277
Seq. No.
                  15620 1.R1040
Contig ID
5'-most EST
                  rca700998350.h1
Method
                  BLASTX
NCBI GI
                  g4191782
BLAST score
                  548
E value
                  2.0e-56
Match length
                  124
% identity
                  86
                  (AC005917) WD-40 repeat protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   9278
Contig ID
                   15621 1.R1040
5'-most EST
                  LIB3072-053-Q1-E1-D8
Seq. No.
                  9279
                  15623 1.R1040
Contig ID
5'-most EST
                  wrq700791003.hl
Method
                  BLASTX
NCBI GI
                   g3461884
BLAST score
                  340
E value
                  1.0e-31
Match length
                   95
% identity
                  72
NCBI Description
                   (AB006082) phosphoribosyl-ATP pyrophosphohydrolase
                   [Arabidopsis thaliana] >gi 3461886 dbj BAA32529 (AB006083)
                  phosphoribosyl-ATP pyrophosphohydrolase [Arabidopsis
                   thaliana]
                   9280
Seq. No.
                  15626 1.R1040
Contig ID
5'-most EST
                  LIB3170-087-Q1-K1-C5
Method
                  BLASTX
NCBI GI
                  g2706450
```

BLAST score 903 2.0e-97 E value Match length 216 % identity 81

(AJ225172) magnesium dependent soluble inorganic NCBI Description

pyrophosphatase [Solanum tuberosum]

Seq. No. 9281



Contig ID 15626_2.R1040 5'-most EST leu701146961.h1

Method BLASTX
NCBI GI g3695383
BLAST score 581
E value 6.0e-60
Match length 124
% identity 86

NCBI Description (AF096370) similar to inorganic pyrophosphatase (Pfam: PF00719 Pyrophosphatase, E-value: 2.7e-88) [Arabidopsis

thaliana]

Seq. No. 9282

Contig ID 15626 3.R1040

5'-most EST uC-gmrominsoy208e01b1

Method BLASTX
NCBI GI g2706450
BLAST score 322
E value 1.0e-29
Match length 75
% identity 79

NCBI Description (AJ225172) magnesium dependent soluble inorganic

pyrophosphatase [Solanum tuberosum]

Seq. No. 9283

Contig ID 15626 4.R1040

5'-most EST jC-gmro02800028b03d1

Method BLASTX
NCBI GI g2706450
BLAST score 257
E value 4.0e-22
Match length 59

Match length 59 % identity 85

NCBI Description (AJ225172) magnesium dependent soluble inorganic

pyrophosphatase [Solanum tuberosum]

Seq. No. 9284

Contig ID 15626_5.R1040

5'-most EST LIB3107-052-Q1-K1-E6

Method BLASTX
NCBI GI g2706450
BLAST score 305
E value 9.0e-28
Match length 85
% identity 71

NCBI Description (AJ225172) magnesium dependent soluble inorganic

pyrophosphatase [Solanum tuberosum]

Seq. No. 9285

Contig ID 15626_6.R1040 5'-most EST kl1701207561.h1

Method BLASTX
NCBI GI g3695383
BLAST score 328
E value 2.0e-30
Match length 70
% identity 86



NCBI Description (AF096370) similar to inorganic pyrophosphatase (Pfam: PF00719 Pyrophosphatase, E-value: 2.7e-88) [Arabidopsis thaliana]

Seq. No. 9286

Contig ID 15627_1.R1040

5'-most EST jC-gmro02800032f02a1

Method BLASTX
NCBI GI g3775999
BLAST score 472
E value 6.0e-47
Match length 102
% identity 86

NCBI Description (AJ010463) RNA helicase [Arabidopsis thaliana]

Seq. No. 9287

Contig ID 15632_1.R1040

5'-most EST jC-gmle01810091a04d1

Method BLASTX
NCBI GI g3702321
BLAST score 541
E value 3.0e-55
Match length 114

Match length 114 % identity 88

NCBI Description (AC005397) putative TGF-beta receptor interacting protein

[Arabidopsis thaliana]

Seq. No. 9288

Contig ID 15632 2.R1040

5'-most EST jC-gmro02910046d09d1

Method BLASTX
NCBI GI g3702321
BLAST score 453
E value 3.0e-45
Match length 91
% identity 93

NCBI Description (AC005397) putative TGF-beta receptor interacting protein

[Arabidopsis thaliana]

Seq. No. 9289

Contig ID 15635_1.R1040 5'-most EST leu701145544.h1

Method BLASTX
NCBI GI g3063698
BLAST score 551
E value 2.0e-56
Match length 179
% identity 61

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 9290

Contig ID 15642_1.R1040

5'-most EST LIB3049-045-Q1-E1-A3

Seq. No. 9291

Contig ID 15645_1.R1040

5'-most EST LIB3138-073-P1-N1-C9



Contig ID 15648 1.R1040

5'-most EST LIB3170-036-Q1-K2-H3

Seq. No. 9293

Contig ID 15648_3.R1040 5'-most EST bth700846272.h1

Seq. No. 9294

Contig ID 15649_1.R1040

5'-most EST jC-qmro02910016h09a1

Seq. No. 9295

Contig ID 15650_1.R1040

5'-most EST LIB3170-034-Q1-K1-B3

Method BLASTX
NCBI GI g3876622
BLAST score 243
E value 2.0e-20
Match length 96

Match length 96 % identity 45

NCBI Description (Z81523) Similarity to Yeast SOH-1 protein (SW:P38633)

[Caenorhabditis elegans]

Seq. No. 9296

Contig ID 15652 1.R1040

5'-most EST LIB3049-040-Q1-E1-B1

Seg. No. 9297

Contig ID 15654 1.R1040

5'-most EST LIB3170-035-Q1-K1-A3

Seq. No. 9298

Contig ID 15659 1.R1040 5'-most EST leu701150418.h1

Method BLASTX
NCBI GI g2982698
BLAST score 290
E value 2.0e-25
Match length 208
% identity 37

NCBI Description (AB006153) tFZR1 [Oncorhynchus mykiss]

Seq. No. 9299

Contig ID 15661 1.R1040

5'-most EST LIB3170-036-Q1-K2-E3

Seq. No. 9300

Contig ID 15662_1.R1040

5'-most EST LIB3049-044-Q1-E1-G8

Method BLASTX
NCBI GI g3341443
BLAST score 333
E value 4.0e-31
Match length 120
% identity 55



```
NCBI Description (AJ223074) acid phosphatase [Glycine max]
                   9301
Seq. No.
                   15665 1.R1040
Contig ID
                   LIB3049-044-Q1-E1-H11
5'-most EST
                   9302
Seq. No.
                   15667 1.R1040
Contig ID
                   dpv70\overline{1}097441.h1
5'-most EST
                   BLASTX
Method
                   q3668089
NCBI GI
                   840
BLAST score
                   6.0e-90
E value
                   298
Match length
                   52
% identity
                   (AC004667) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   15667 2.R1040
Contig ID
                   ncj70\overline{0}982585.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3668089
                   162
BLAST score
                   4.0e-11
E value
                   55
Match length
                   51
% identity
                   (AC004667) unknown protein [Arabidopsis thaliana]
NCBI Description
                   9304
Seq. No.
                   15672 1.R1040
Contig ID
                   uC-gmropic096g06b1
5'-most EST
                   BLASTX
Method
NCBI GI
                    g2832632
                    616
BLAST score
                    6.0e-64
E value
                   191
Match length
% identity
                    68
                   (AL021711) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                    15673 1.R1040
Contig ID
                    ckk70\overline{0}605877.h2
5'-most EST
                    9306
Seq. No.
                    15678 1.R1040
Contig ID
                    LIB3049-044-Q1-E1-F11
5'-most EST
                    BLASTX
Method
                    g3560243
NCBI GI
BLAST score
                    169
                    8.0e-12
E value
                    74
Match length
% identity
                    (AL031532) putative hydrolase [Schizosaccharomyces pombe]
NCBI Description
Seq. No.
                    9307
```

1656

15678 2.R1040

LIB3106-035-Q1-K1-G12

Contig ID

5'-most EST



```
BLASTX
Method
                  g3560243
NCBI GI
                  192
BLAST score
                  3.0e-14
E value
                  99
Match length
                  56
% identity
NCBI Description (AL031532) putative hydrolase [Schizosaccharomyces pombe]
Seq. No.
                  9308
                  15679 1.R1040
Contig ID
                  LIB3170-036-Q1-K1-D3
5'-most EST
                  BLASTX
Method
                  q3319357
NCBI GI
                   191
BLAST score
                   2.0e-14
E value
                   67
Match length
                   57
% identity
                   (AF077407) contains similarity to phosphoenolpyruvate
NCBI Description
                   synthase (ppsA) (GB:AE001056) [Arabidopsis thaliana]
                   9309
Seq. No.
                   15685 1.R1040
Contig ID
                   qsv701049807.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3687224
                   378
BLAST score
                   4.0e-36
E value
                   126
Match length
                   59
% identity
                   (AC005169) putative N-acetyl-gamma-glutamyl-phosphate
NCBI Description
                   reductase [Arabidopsis thaliana]
                   9310
Seq. No.
                   15685 2.R1040
Contig ID
                   LIB3109-017-Q1-K2-F7
5'-most EST
                   BLASTX
Method
                   g3687224
NCBI GI
BLAST score
                   208
                   2.0e-16
E value
                   71
Match length
                   58
% identity
                   (AC005169) putative N-acetyl-gamma-glutamyl-phosphate
NCBI Description
                   reductase [Arabidopsis thaliana]
Seq. No.
                   9311
                   15689 1.R1040
Contig ID
                   LIB3170-033-Q1-J1-E3
5'-most EST
                   9312
Seq. No.
                   15690 1.R1040
Contig ID
                   LIB3049-008-Q1-E1-B4
 5'-most EST
Method
                   BLASTN
                   g4164411
NCBI GI
BLAST score
                   34
E value
                   2.0e-09
Match length
                   38
```

97

% identity



NCBI Description Pisum sativum mitochondrial rp15, rps14 and cob genes

Seq. No. 9313

Contig ID 15692_1.R1040 5'-most EST kl1701202946.h1

Method BLASTX
NCBI GI g3063695
BLAST score 321
E value 2.0e-29
Match length 109
% identity 62

NCBI Description (AL022537) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9314

Contig ID 15701 1.R1040

5'-most EST LIB3049-044-Q1-E1-E12

Method BLASTX
NCBI GI g2346988
BLAST score 174
E value 3.0e-12
Match length 98
% identity 25

NCBI Description (AB006606) ZPT4-4 [Petunia x hybrida]

Seq. No. 9315

Contig ID 15704 1.R1040

5'-most EST uC-gmrominsoy318b03b1

Method BLASTN
NCBI GI g11303
BLAST score 637
E value 0.0e+00
Match length 725
% identity 97

NCBI Description G.max chloroplast mRNA for ndhK polypeptide and orf 158/159

Seq. No. 9316

Contig ID 15704_2.R1040

5'-most EST LIB3109-031-Q1-K1-E4

Method BLASTN
NCBI GI g881440
BLAST score 232
E value 1.0e-127
Match length 328
% identity 93

NCBI Description Lupinus luteus NADH-plastoquinone oxidoreductase (ndhC)

gene, chloroplast gene encoding chloroplast protein,

complete cds

Seq. No. 9317

Contig ID 15706 1.R1040 5'-most EST hrw701058184.h1

Method BLASTX
NCBI GI g3320120
BLAST score 162
E value 6.0e-11
Match length 65
% identity 48



NCBI Description (U66669) 3-hydroxyisobutyryl-coenzyme A hydrolase [Homo sapiens]

Seq. No.

9318

BLASTX

205

q3935166

6.0e-16

Contig ID 5'-most EST 15711_1.R1040 hyd700728580.h1

Method NCBI GI BLAST score E value

Match length 141 % identity 38

NCBI Description (AC004557) F17L21.9 [Arabidopsis thaliana]

Seq. No.

9319

Contig ID

15712_1.R1040

5'-most EST

LIB3106-022-Q1-K1-B9

Method BLASTX
NCBI GI g2500116
BLAST score 1786
E value 0.0e+00
Match length 402
% identity 87

NCBI Description GLUTATHIONE REDUCTASE, CYTOSOLIC (GR) (GRASE) (GOR2)

>gi 1370285 emb CAA66924 (X98274) glutathione reductase

[Pisum sativum]

Seq. No.

Contig ID 5'-most EST

15712 3.R1040

nsy700645585.h1

Seq. No.

9321

9320

Contig ID 5'-most EST

15715_1.R1040 kl1701214315.h1

Seq. No.

9322

Contig ID

15716 1.R1040

5'-most EST

LIB3049-044-Q1-E1-C2

Seq. No.

9323

Contig ID

15720 1.R1040

5'-most EST

LIB3170-035-Q1-K1-E2

Method BLASTX
NCBI GI g2988398
BLAST score 160
E value 1.0e-10
Match length 108
% identity 37

NCBI Description (AC004381) Unknown gene product [Homo sapiens]

Seq. No.

9324

Contig ID 15720 2.R1040

5'-most EST

LIB3049-033-Q1-E1-C11

Seq. No.

9325

Contig ID

15720 3.R1040

5'-most EST

LIB3106-065-P1-K1-E4



```
9326
Seq. No.
                  15721 1.R1040
Contig ID
                  LIB3049-044-Q1-E1-C8
5'-most EST
                  BLASTX
Method
                  q3297808
NCBI GI
                  210
BLAST score
                  2.0e-33
E value
                  136
Match length
% identity
NCBI Description (AL031032) putative protein [Arabidopsis thaliana]
                  9327
Seq. No.
                  15727 1.R1040
Contig ID
                  LIB3049-044-Q1-E1-A3
5'-most EST
                   9328
Seq. No.
                   15738 1.R1040
Contig ID
                   sat701004184.h1
5'-most EST
                  BLASTX
Method
                   g4263711
NCBI GI
                   658
BLAST score
                   9.0e-69
E value
                   250
Match length
                   52
% identity
                   (AC006223) putative CCR4-associated transcription factor
NCBI Description
                   [Arabidopsis thaliana]
                   9329
Seq. No.
                   15738 2.R1040
Contig ID
                   LIB3092-037-Q1-K1-B3
5'-most EST
                   BLASTX
Method
                   q4263711
NCBI GI
                   285
BLAST score
                   2.0e-25
E value
                   121
Match length
% identity
                   47
                   (AC006223) putative CCR4-associated transcription factor
NCBI Description
                   [Arabidopsis thaliana]
                   9330
Seq. No.
                   15740 1.R1040
Contig ID
                   LIB3049-044-Q1-E1-B6
5'-most EST
Method
                   BLASTX
                   g128378
NCBI GI
BLAST score
                   205
                   3.0e-16
E value
                   76
Match length
 % identity
                   50
                   NONSPECIFIC LIPID-TRANSFER PROTEIN A (NS-LTP A)
NCBI Description
                   (PHOSPHOLIPID TRANSFER PROTEIN) (PLTP)
                   >gi_82031_pir__S07142 nonspecific lipid transfer protein -
                   castor bean >gi 224909 prf 1204170A protein, nonspecific
```

Contig ID 15742_1.R1040

lipid transfer [Ricinus communis]



5'-most EST dkc700968011.h1

Method BLASTX
NCBI GI g2245138
BLAST score 858
E value 4.0e-92
Match length 258
% identity 70

NCBI Description (Z97344) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9332

Contig ID 15742_2.R1040

5'-most EST jC-gmst02400025c09a1

Method BLASTX
NCBI GI g2245138
BLAST score 374
E value 8.0e-36
Match length 121
% identity 64

NCBI Description (Z97344) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9333

Contig ID 15749 1.R1040

5'-most EST LIB3049-043-Q1-E1-H1

Seq. No. 9334

Contig ID 15754_1.R1040 5'-most EST wvk700682131.h1

Method BLASTX
NCBI GI g3668080
BLAST score 912
E value 2.0e-98
Match length 289
% identity 67

NCBI Description (AC004667) unknown protein [Arabidopsis thaliana]

Seq. No. 9335

Contig ID 15754_2.R1040

5'-most EST LIB3106-063-Q1-K1-C9

Seq. No. 9336

Contig ID 15757_1.R1040

5'-most EST LIB3170-035-Q1-K1-D2

Seq. No. 9337

Contig ID 15759 1.R1040

5'-most EST LIB3093-044-Q1-K1-F3

Method BLASTX
NCBI GI 94091117
BLAST score 535
E value 2.0e-54
Match length 171
% identity 62

NCBI Description (AF047428) nucleic acid binding protein [Oryza sativa]

Seq. No. 9338

Contig ID 15759 2.R1040

5'-most EST LIB3049-043-Q1-E1-G2



```
BLASTX
Method
                  g166410
NCBI GI
                  219
BLAST score
                  6.0e-18
E value
                  48
Match length
% identity
NCBI Description (L07291) Alfin-1 [Medicago sativa]
                   9339
Seq. No.
                   15761 1.R1040
Contig ID
                  LIB3170-036-Q1-K1-H2
5'-most EST
                  BLASTX
Method
                   g3201554
NCBI GI
                   767
BLAST score
                   1.0e-81
E value
                   219
Match length
                   68
% identity
                   (AJ006501) beta-D-glucosidase [Tropaeolum majus]
NCBI Description
Seq. No.
                   15763 1.R1040
Contig ID
                   jC-gmst02400001d02a1
5'-most EST
                   BLASTX
Method
                   g2190007
NCBI GI
                   623
BLAST score
                   2.0e-65
E value
                   307
Match length
                   43
% identity
                   (AB004109) phosphatidylserine synthase II [Cricetulus
NCBI Description
                   griseus]
                   9341
Seq. No.
                   15764 1.R1040
Contig ID
                   LIB3170-036-Q1-K1-B2
5'-most EST
                   9342
Seq. No.
                   15769 1.R1040
Contig ID
                   jsh701063734.h1
5'-most EST
                   BLASTX
Method
                   g1076670
NCBI GI
BLAST score
                   483
                   1.0e-48
E value
                   99
Match length
                   91
% identity
                   NADH dehydrogenase (EC 1.6.99.3) - potato
NCBI Description
                   >gi 668985 emb_CAA59062_ (X84319) NADH dehydrogenase
                   [Solanum tuberosum]
                    9343
 Seq. No.
                   15770 1.R1040
Contig ID
                   uC-gmrominsoy246d02b1
 5'-most EST
Method
                   BLASTX
                   q2059326
NCBI GI
 BLAST score
                   181
E value
                    4.0e-13
                    109
Match length
```

35

% identity



NCBI Description (D67067) thymic epithelial cell surface antigen [Mus musculus]

Seq. No. 9344

Contig ID 15770_2.R1040

5'-most EST LIB3056-007-Q1-N1-F9

0245

Seq. No. Contig ID

5'-most EST

15771_1.R1040 kmv700739716.h1

Seq. No. 9346

Contig ID 15775_1.R1040

5'-most EST LIB3170-036-Q1-K2-D2

Method BLASTX
NCBI GI g4467126
BLAST score 393
E value 8.0e-38
Match length 175

% identity 46

NCBI Description (AL035538) guanine nucleotide-exchange protein-like

[Arabidopsis thaliana]

Seq. No. 9347

Contig ID 15777_1.R1040

5'-most EST LIB3049-043-Q1-E1-G11

Method BLASTX
NCBI GI 94335745
BLAST score 583
E value 8.0e-60
Match length 311
% identity 42

NCBI Description (AC006284) putative hydrolase (contains an

esterase/lipase/thioesterase active site serine domain

(prosite: PS50187) [Arabidopsis thaliana]

Seq. No. 9348

Contig ID 15781_1.R1040

5'-most EST LIB3074-040-Q1-K1-A7

Seq. No. 9349

Contig ID 15781_3.R1040 5'-most EST gsv701054002.h1

Seq. No. 9350

Contig ID 15782_1.R1040 5'-most EST vwf700674333.h1

Method BLASTX
NCBI GI g1922242
BLAST score 272
E value 9.0e-24
Match length 67
% identity 85

NCBI Description (Y10084) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9351

Contig ID 15792_1.R1040



```
LIB3049-043-Q1-E1-E12
5'-most EST
                  9352
Seq. No.
                  15795 1.R1040
Contig ID
                  LIB3138-054-Q1-N1-C2
5'-most EST
                  BLASTX
Method
                  g2980762
NCBI GI
BLAST score
                  160
                  8.0e-11
E value
Match length
                  68
% identity
NCBI Description (AL022198) hypothetical protein [Arabidopsis thaliana]
                   9353
Seq. No.
                   15799 1.R1040
Contig ID
                  LIB3049-043-Q1-E1-B4
5'-most EST
                  BLASTX
Method
                   g3540181
NCBI GI
                   521
BLAST score
                   7.0e-53
E value
                   227
Match length
                   50
% identity
NCBI Description (AC004122) Unknown protein [Arabidopsis thaliana]
                   9354
Seq. No.
                   15807 1.R1040
Contig ID
                   LIB3049-043-Q1-E1-C3
5'-most EST
                   9355
Seq. No.
                   15808 1.R1040
Contig ID
                   LIB3170-034-Q1-K1-E1
5'-most EST
                   9356
Seq. No.
                   15812 1.R1040
Contig ID
                   LIB3049-043-Q1-E1-C7
5'-most EST
                   9357
Seq. No.
                   15817 1.R1040
Contig ID
                   LIB3049-042-Q1-E1-H6
5'-most EST
                   9358
Seq. No.
                   15819 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220093d01a1
Seq. No.
                   9359
                   15820 1.R1040
Contig ID
                   LIB3049-042-Q1-E1-H9
 5'-most EST
                   BLASTX
Method
NCBI GI
                   g133872
                   445
BLAST score
                   4.0e-44
E value
                   115
Match length
                   73
 % identity
                   30S RIBOSOMAL PROTEIN S1, CHLOROPLAST PRECURSOR (CS1)
 NCBI Description
                   >gi_282838_pir__S26494 ribosomal protein S1, chloroplast -
                    spinach >gi_322404_pir__A44121 small subunit ribosomal
```

protein CS1, CS-S2 - spinach >gi_18060_emb_CAA46927_



(X66135) ribosomal protein S1 [Spinacia oleracea] >gi 170143 (M82923) chloroplast ribosomal protein S1 [Spinacia oleracea]

9360 Seq. No.

15827 1.R1040 Contig ID

5'-most EST LIB3139-041-P1-N1-B4

Seq. No.

9361

Contig ID

15840 1.R1040 LIB3049-042-Q1-E1-G4

5'-most EST Method BLASTX NCBI GI BLAST score

q3738289 416 1.0e-40

E value Match length % identity

116 69

NCBI Description

(AC005309) Not56-like protein [Arabidopsis thaliana]

Seq. No.

9362

Contig ID

15849 1.R1040

5'-most EST

LIB3049-042-Q1-E1-E2

Seq. No.

9363

Contig ID

15857 1.R1040

5'-most EST

LIB3049-042-Q1-E1-F10

Method NCBI GI BLAST score

q3894191 194

BLASTX

E value Match length

8.0e-15 102 41

% identity NCBI Description

(AC005662) unknown protein [Arabidopsis thaliana]

Seq. No.

9364

Contig ID

15858 1.R1040

5'-most EST

LIB3049-042-Q1-E1-F11

Seq. No.

9365

Contig ID

15872 1.R1040

5'-most EST

LIB3139-004-P1-N1-D1

Seq. No.

9366

Contig ID

15873 1.R1040

5'-most EST

LIB3170-030-Q1-K1-F12

Method NCBI GI BLAST score BLASTX g2832629 321

E value

8.0e-30

Match length % identity

127

NCBI Description

46 (AL021711) 4-coumarate-CoA ligase - like [Arabidopsis

thaliana]

Seq. No.

9367

Contig ID 5'-most EST 15875 1.R1040

uC-gmflminsoy031a04b1



```
BLASTX
Method
                  g3309086
NCBI GI
                  1046
BLAST score
                  1.0e-114
E value
                  226
Match length
                   90
% identity
                   (AF076253) calcineurin B-like protein 3 [Arabidopsis
NCBI Description
                   thaliana]
                   9368
Seq. No.
                   15875 2.R1040
Contig ID
                   LIB3049-042-Q1-E1-D7
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3309084
                   158
BLAST score
                   1.0e-10
E value
Match length
                   48
                   65
% identity
                   (AF076252) calcineurin B-like protein 2 [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   9369
                   15881 1.R1040
Contig ID
                   leu701150201.h1
5'-most EST
Method
                   BLASTX
                   q4097547
NCBI GI
BLAST score
                   549
                   4.0e-56
E value
Match length
                   167
% identity
                   66
NCBI Description (U64906) ATFP3 [Arabidopsis thaliana]
                   9370
Seq. No.
                   15881 2.R1040
Contig ID
                   LIB3072-057-Q1-K1-H7
5'-most EST
                   BLASTX
Method
                   q4097547
NCBI GI
BLAST score
                   291
                   4.0e-26
E value
                   66
Match length
                   41
% identity
NCBI Description (U64906) ATFP3 [Arabidopsis thaliana]
                   9371
Seq. No.
                   15881 3.R1040
Contig ID
                   LIB3170-083-Q1-K1-D11
5'-most EST
                   BLASTX
Method
                   g4097547
NCBI GI
BLAST score
                   214
                   3.0e-17
E value
                   53
Match length
% identity
                   39
                  (U64906) ATFP3 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   9372
```

1666

15891 1.R1040

 $asn70\overline{1}136815.h1$

Contig ID

5'-most EST



Method BLASTX NCBI GI g2505874 BLAST score 1051 1.0e-119 E value Match length 250 % identity 80

(Y12776) putative kinase [Arabidopsis thaliana] NCBI Description

Seq. No. Contig ID 9373 15894 1.R1040

5'-most EST

LIB3073-014-Q1-K1-E8

Seq. No. Contig ID

9374 15894 2.R1040 5'-most EST jsh701067288.h1

Seq. No.

9375

15894 3.R1040 Contig ID

5'-most EST

LIB3073-013-Q1-K1-E8

Seq. No. Contig ID 9376

15894 4.R1040

9377

5'-most EST

jC-gmst02400008a10a1

Seq. No. Contig ID 5'-most EST

15894 5.R1040 $seb70\overline{0}651282.h1$

Seq. No.

9378

Contig ID 5'-most EST 15898 1.R1040 pxt700944642.hl

Method BLASTX NCBI GI q3540197 BLAST score 211 E value 1.0e-16 72

Match length % identity

56

(AC004260) Unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

9379

Contig ID

15900 1.R1040

5'-most EST

LIB3170-024-Q1-K1-E1

Method BLASTX NCBI GI q4336747 BLAST score 416 2.0e-40 E value Match length 281

% identity

36 NCBI Description

(AF104924) unconventional myosin heavy chain [Zea mays]

Seq. No.

9380

Contig ID

15901 1.R1040

5'-most EST

LIB3049-041-Q1-E1-H5

Method BLASTX NCBI GI g2245128 BLAST score 438 E value 3.0e-43



Match length 126 % identity 65

NCBI Description (Z97344) peroxidase [Arabidopsis thaliana]

Seq. No.

9381

Contig ID 5'-most EST

15902_1.R1040 hyd700727214.h1

Seq. No.

9382

Contig ID 5'-most EST 15902_2.R1040 zsg701120876.h1

Method BLASTX
NCBI GI g2351846
BLAST score 269
E value 2.0e-23
Match length 111
% identity 52

NCBI Description (U93050) poly(A) binding protein II [Mus musculus]

Seq. No.

9383

Contig ID 5'-most EST

15910 1.R1040

uC-gmropic044c11b1

Method BLASTX
NCBI GI g4415942
BLAST score 240
E value 2.0e-35
Match length 161
% identity 53

NCBI Description (AC006418) hypothetical protein [Arabidopsis thaliana]

Seq. No.

9384

Contig ID

15910 2.R1040

5'-most EST

LIB3139-050-P1-N1-B6

Method BLASTX
NCBI GI g4415942
BLAST score 226
E value 3.0e-18
Match length 118
% identity 39

NCBI Description (AC006418) hypothetical protein [Arabidopsis thaliana]

Seq. No.

9385

Contig ID . 5 - most EST

15913_1.R1040 kl1701211472.h1

Method BLASTX
NCBI GI g4432855
BLAST score 1127
E value 1.0e-123
Match length 339
% identity 71

NCBI Description (AC006300) unknown protein [Arabidopsis thaliana]

Seq. No.

9386

Contig ID

15915_1.R1040

5'-most EST

LIB3170-041-Q1-K1-B3

Method BLASTN NCBI GI q3046856

1668



```
BLAST score
                  4.0e-23
E value
                  141
Match length
                  86
% identity
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MXI22, complete sequence [Arabidopsis thaliana]
Seq. No.
                  9387
                  15917 1.R1040
Contig ID
                  LIB3049-040-Q1-E1-H6
5'-most EST
                  BLASTX
Method
                  q1362078
NCBI GI
                  578
BLAST score
                  2.0e-59
E value
                  216
Match length
                   60
% identity
                  endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG1)
NCBI Description
                   - common nasturtium >gi_311835_emb_CAA48324_ (X68254)
                   cellulase [Tropaeolum majus]
                   9388
Seq. No.
                   15921 1.R1040
Contig ID
                  LIB3138-023-Q1-N1-G4
5'-most EST
Method
                   BLASTX
NCBI GI
                   q2920839
                   337
BLAST score
                   2.0e-31
E value
                   129
Match length
% identity
                   51
                   (U95136) Os-FIERG2 gene product [Oryza sativa]
NCBI Description
                   9389
Seq. No.
                   15924 1.R1040
Contig ID
                   LIB3093-024-Q1-K1-A6
5'-most EST
Method
                   BLASTX
                   g3063709
NCBI GI
                   358
BLAST score
                   5.0e-34
E value
                   135
Match length
% identity
                  (AL022537) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   9390
                   15926 1.R1040
Contig ID
                   LIB3170-022-Q1-K1-E6
5'-most EST
                   BLASTN
Method
```

Method BLASTN
NCBI GI g3985958
BLAST score 41
E value 2.0e-13
Match length 290
% identity 82

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MZN1, complete sequence [Arabidopsis thaliana]

Seq. No. 9391

Contig ID 15926 3.R1040

5'-most EST LIB3170-032-Q1-K1-E4



Method BLASTN
NCBI GI g3985958
BLAST score 33
E value 5.0e-09
Match length 182
% identity 82

% identity 82 NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MZN1, complete sequence [Arabidopsis thaliana]

Seq. No. 9392

Contig ID 15931_1.R1040

5'-most EST LIB3049-041-Q1-E1-G9

Method BLASTX
NCBI GI g3914156
BLAST score 357
E value 6.0e-34
Match length 94
% identity 70

NCBI Description NUCLEOLAR PROTEIN AT BAND 60B >gi_2661227 (AF017230)

nucleolar protein at band 60B [Drosophila melanogaster] >gi_4322557_gb_AAD16092_ (AF089837) minifly protein

[Drosophila melanogaster]

Seq. No. 9393

Contig ID 15939 1.R1040

5'-most EST LIB3073-013-Q1-K1-C10

Method BLASTX
NCBI GI 94559358
BLAST score 422
E value 2.0e-41
Match length 100
% identity 77

NCBI Description (AC006585) putative steroid binding protein [Arabidopsis

thaliana]

Seq. No. 9394

Contig ID 15939_2.R1040

5'-most EST LIB3049-041-Q1-E1-E11

Method BLASTX
NCBI GI 94559358
BLAST score 398
E value 1.0e-38
Match length 99
% identity 73

NCBI Description (AC006585) putative steroid binding protein [Arabidopsis

thaliana]

Seq. No. 9395

Contig ID 15941_1.R1040 5'-most EST gbt700547911.h1

Seq. No. 9396

Contig ID 15942_1.R1040

5'-most EST LIB3049-041-Q1-E1-E3

Seq. No. 9397

Contig ID 15950 1.R1040



```
taw700656026.h1
5'-most EST
Method
                  BLASTX
                  g2244822
NCBI GI
                  235
BLAST score
                  2.0e-19
E value
Match length
                   63
                   70
% identity
                   (Z97336) RNA polymerase II fifth largest subunit homolog
NCBI Description
                   [Arabidopsis thaliana]
                   9398
Seq. No.
                   15952 1.R1040
Contig ID
                  LIB3049-016-Q1-E1-C3
5'-most EST
                  BLASTX
Method
NCBI GI
                   q2896687
BLAST score
                   165
E value
                   2.0e-11
Match length
                   87
                   48
% identity
                   (AL021897) hypothetical protein Rv1050 [Mycobacterium
NCBI Description
                   tuberculosis]
Seq. No.
                   9399
                   15952 2.R1040
Contig ID
5'-most EST
                   fde700876935.h1
                   9400
Seq. No.
                   15953 1.R1040
Contig ID
                   jC-gmro02910005e12d1
5'-most EST
Method
                   BLASTX
                   g3033400
NCBI GI
                   245
BLAST score
                   4.0e-20
E value
Match length
                   104
% identity
                   50
                   (AC004238) putative Ser/Thr protein kinase [Arabidopsis
NCBI Description
                   thaliana]
                   9401
Seq. No.
                   15955 1.R1040
Contig ID
                   fua701038882.h1
5'-most EST
                   BLASTX
Method
                   g2827699
NCBI GI
                   422
BLAST score
                   3.0e-41
E value
                   206
Match length
                   52
% identity
                   (AL021684) predicted protein [Arabidopsis thaliana]
NCBI Description
                   9402
Seq. No.
                   15955 2.R1040
Contig ID
                   crh700853141.h1
5'-most EST
                   BLASTX
```

Method BLASTX
NCBI GI g2827699
BLAST score 120
E value 7.0e-14
Match length 106



% identity 52
NCBI Description (AL021684) predicted protein [Arabidopsis thaliana]

Seq. No. 9403

Contig ID 15957_1.R1040 5'-most EST zhf700951979.h1

Seq. No. 9404

Contig ID 15966_1.R1040

5'-most EST LIB3049-041-Q1-E1-C5

Seq. No. 9405

Contig ID 15971_1.R1040 5'-most EST q4294765

Seq. No. 9406

Contig ID 15972 1.R1040

5'-most EST LIB3072-061-Q1-K1-H6

Seq. No. 9407

Contig ID 15974_1.R1040

5'-most EST LIB3049-040-Q1-E1-H7

Seq. No. 9408

Contig ID 15976_1.R1040 5'-most EST gsv701052356.h1

Method BLASTX
NCBI GI g1170606
BLAST score 673
E value 1.0e-70
Match length 188

Match length 188 % identity 70

NCBI Description ADENYLATE KINASE, CHLOROPLAST (ATP-AMP TRANSPHOSPHORYLASE)

>gi_629863_pir__S45634 adenylate kinase (EC 2.7.4.3),
chloroplast - maize >gi 3114421_pdb_1ZAK_A Chain A,

Adenylate Kinase From Maize In Complex With The Inhibitor

P1, P5-Bis (Adenosine-5'-) pentaphosphate (Ap5a)

>gi 3114422 pdb 1ZAK B Chain B, Adenylate Kinase From Maize

In Complex With The Inhibitor

P1, P5-Bis (Adenosine-5'-) pentaphosphate (Ap5a)

Seq. No. 9409

Contig ID 15986_1.R1040

5'-most EST LIB3049-041-Q1-E1-B10

Seq. No. 9410

Contig ID 15991 1.R1040

5'-most EST LIB3107-012-Q1-K1-G6

Seq. No. 9411

Contig ID 15991 2.R1040

5'-most EST LIB3049-040-Q1-E1-G1

Seq. No. 9412

Contig ID 15994 1.R1040

5'-most EST LIB3049-040-Q1-E1-G12



Contig ID 15995 1.R1040 5'-most EST nsy700645633.h1

Method BLASTN
NCBI GI g3868852
BLAST score 268
E value 1.0e-149
Match length 588
% identity 88

NCBI Description Vigna radiata ARG8 mRNA for GPI-anchored protein, complete

cds

Seq. No. 9414

Contig ID 15995_2.R1040 5'-most EST ssr700557224.h1

Seq. No. 9415

Contig ID 15995_3.R1040

5'-most EST LIB3049-040-Q1-E1-G2

Method BLASTN
NCBI GI g3868852
BLAST score 103
E value 1.0e-50
Match length 333
% identity 84

NCBI Description Vigna radiata ARG8 mRNA for GPI-anchored protein, complete

cds

Seq. No. 9416

Contig ID 15995 4.R1040

5'-most EST LIB3093-018-Q1-K1-H7

Method BLASTN
NCBI GI g3868852
BLAST score 155
E value 1.0e-81
Match length 293
% identity 89

NCBI Description Vigna radiata ARG8 mRNA for GPI-anchored protein, complete

cds

Seq. No. 9417

Contig ID 15996_1.R1040 5'-most EST ncj700988212.h1

Method BLASTX
NCBI GI g2655037
BLAST score 274
E value 4.0e-36
Match length 261
% identity 35

NCBI Description (AF019952) tumor suppressing STF cDNA 1 [Homo sapiens]

>gi 4507703_ref_NP_003301.1_pTSSC1_ tumor suppressing

subtransferable candidate

Seq. No. 9418

Contig ID 15997_1.R1040 5'-most EST uxk700671968.h1



Contig ID 15997_2.R1040

5'-most EST LIB3049-040-Q1-E1-G4

Seq. No. 9420

Contig ID 16005_1.R1040

5'-most EST LIB3170-030-Q1-K1-D11

Seq. No. 9421

Contig ID 16007 1.R1040

5'-most EST LIB3049-008-Q1-E1-B7

Method BLASTX
NCBI GI 94309700
BLAST score 257
E value 5.0e-22
Match length 94
% identity 53

NCBI Description (AC006266) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9422

Contig ID 16009_1.R1040

5'-most EST LIB3056-013-Q1-N1-G8

Seq. No. 9423

Contig ID 16009 2.R1040

5'-most EST LIB3049-040-Q1-E1-E2

Seq. No. 9424

Contig ID 16009_3.R1040

5'-most EST LIB3138-019-Q1-N1-D1

Seq. No. 9425

Contig ID 16013_1.R1040

5'-most EST LIB3074-040-Q1-K1-C4

Method BLASTX
NCBI GI g3283057
BLAST score 237
E value 1.0e-19
Match length 87
% identity 55

NCBI Description (AF054617) one helix protein [Arabidopsis thaliana]

Seq. No.

Contig ID 16016_1.R1040 5'-most EST zzp700834673.h1

Seq. No. 9427

Contig ID 16017 1.R1040

5'-most EST LIB $30\overline{4}$ 9-040-Q1-E1-F12

9426

Seq. No. 9428

Contig ID 16021 1.R1040

5'-most EST LIB3049-040-Q1-E1-F5

Seq. No. 9429

Contig ID 16025 1.R1040

5'-most EST LIB3049-040-Q1-E1-C6



```
Method
                   BLASTX
NCBI GI
                   g3068713
BLAST score
                   243
                   1.0e-20
E value
                   67
Match length
                   75
% identity
NCBI Description (AF049236) unknown [Arabidopsis thaliana]
                   9430
Seq. No.
                   16029 1.R1040
Contiq ID
                   LIB3049-040-Q1-E1-D2
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4559351
                   176
BLAST score
                   1.0e-12
E value
                   45
Match length
                   71
% identity
                  (AC006585) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   16029 2.R1040
Contig ID
                   uaw700663387.h1
5'-most EST
                   BLASTX
Method
                   g4101574
NCBI GI
                   276
BLAST score
                   6.0e-24
E value
                   187
Match length
                   36
% identity
NCBI Description (AF004876) 54TMp [Homo sapiens]
Seq. No.
                   9432
                   16029 3.R1040
Contig ID
5'-most EST
                   dpv70\overline{1}099693.h1
                   BLASTN
Method
                   q4191760
NCBI GI
                   33
BLAST score
                   6.0e-09
E value
Match length
                   155
% identity
                   81
                   Genomic sequence for Arabidopsis thaliana BAC F17F8,
NCBI Description
                   complete sequence [Arabidopsis thaliana]
Seq. No.
                   9433
Contig ID
                   16042 1.R1040
5'-most EST
                   LIB3049-040-Q1-E1-E11
                   BLASTX
Method
NCBI GI
                   g3157923
BLAST score
                   162
E value
                   2.0e-11
Match length
                   81
% identity
                   (AC002131) F12F1.7 [Arabidopsis thaliana]
NCBI Description
```

Contig ID 16049_1.R1040

5'-most EST LIB3049-040-Q1-E1-B5



```
9435
Seq. No.
                  16060 1.R1040
Contig ID
                  kmv70\overline{0}738950.h1
5'-most EST
                  BLASTX
Method
                  g3445209
NCBI GI
                  1036
BLAST score
                  1.0e-113
E value
Match length
                  368
% identity
                  (AC004786) putative serine carboxypeptidase I [Arabidopsis
NCBI Description
                  thaliana]
                   9436
Seq. No.
                   16063 1.R1040
Contig ID
                  uC-gmflminsoy047e02b1
5'-most EST
                   9437
Seq. No.
                   16063 2.R1040
Contig ID
                   LIB3049-055-Q1-E1-C10
5'-most EST
                   9438
Seq. No.
                   16064 1.R1040
Contig ID
                   bth700844340.h1
5'-most EST
                   BLASTX
Method
                   g1914683
NCBI GI
                   422
BLAST score
                   2.0e-41
E value
                   184
Match length
                   57
% identity
NCBI Description (Y12013) RAD23, isoform I [Daucus carota]
                   9439
Seq. No.
                   16064 2.R1040
Contig ID
                   LIB3049-039-Q1-E1-H5
5'-most EST
                   BLASTX
Method
                   q1914683
NCBI GI
                   251
BLAST score
                   1.0e-21
E value
                   88
Match length
                   62
% identity
NCBI Description (Y12013) RAD23, isoform I [Daucus carota]
                   9440
Seq. No.
Contig ID
                   16065 1.R1040
                   LIB3073-017-Q1-K1-E1
5'-most EST
                   BLASTX
Method
                   q2253442
NCBI GI
BLAST score
                   221
E value
                   1.0e-17
                   61
Match length
% identity
                   (AF007784) LTCOR11 [Lavatera thuringiaca]
NCBI Description
Seq. No.
                   9441
```

Contig ID 16065_2.R1040

5'-most EST LIB3051-041-Q1-K1-B4

Method BLASTX

1676



```
q2253442
NCBI GI
                  223
BLAST score
                  8.0e-18
E value
Match length
                  61
% identity
                  62
NCBI Description (AF007784) LTCOR11 [Lavatera thuringiaca]
Seq. No.
                  16068 1.R1040
Contig ID
                  sat701002941.hl
5'-most EST
                   9443
Seq. No.
                   16069 1.R1040
Contig ID
5'-most EST
                  uC-gmronoir060f12b1
                   9444
Seq. No.
                   16070 1.R1040
Contig ID
                   zsg701117861.h2
5'-most EST
                   9445
Seq. No.
                   16074 1.R1040
Contig ID
                   LIB3170-030-Q1-K1-D10
5'-most EST
                   9446
Seq. No.
                   16075 1.R1040
Contig ID
                   LIB3170-031-Q1-K1-G10
5'-most EST
                   9447
Seq. No.
                   16078 1.R1040
Contig ID
                   LIB3139-044-P1-N1-D4
5'-most EST
                   BLASTX
Method
                   g4206789
NCBI GI
BLAST score
                   835
E value
                   1.0e-89
                   253
Match length
                   67
% identity
                   (AF112864) syntaxin-related protein At-SYR1 [Arabidopsis
NCBI Description
                   thaliana]
                   9448
Seq. No.
                   16079 1.R1040
Contig ID
                   LIB3170-035-Q1-K1-E3
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3033395
BLAST score
                   168
                   2.0e-11
E value
                   60
Match length
% identity
```

(ACO04238) putative zinc-finger protein [Arabidopsis NCBI Description

thaliana]

9449 Seq. No.

16080 1.R1040 Contig ID

5'-most EST LIB3049-039-Q1-E1-F9

Seq. No. 9450

Contig ID 16081 1.R1040

1677



5'-most EST LIB3049-039-Q1-E1-G1

Method BLASTX
NCBI GI g3860255
BLAST score 222
E value 6.0e-18
Match length 128
% identity 41

NCBI Description (AC005824) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9451

Contig ID 16083_1.R1040 5'-most EST gsv701054107.h1

Method BLASTN
NCBI GI g3869074
BLAST score 38
E value 1.0e-11
Match length 50
% identity 94

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MMI9, complete sequence [Arabidopsis thaliana]

Seg. No. 9452

Contig ID 16089 1.R1040

5'-most EST LIB3049-039-Q1-E1-G7

Seq. No. 9453

Contig ID 16090 1.R1040

5'-most EST LIB3049-039-Q1-E1-G9

Seq. No. 9454

Contig ID 16093 1.R1040

5'-most EST LIB31 $\overline{3}$ 9-026-P1-N1-H3

Method BLASTX
NCBI GI g1532163
BLAST score 211
E value 2.0e-16
Match length 90
% identity 50

NCBI Description (U63815) similar to glutaredoxin encoded by GenBank

Accession Number Z49699; localized according to blastn similarity to EST sequences; therefore, the coding span corresponds only to an area of similarity since the

initation codon and stop c

Seq. No. 9455

Contig ID 16093_2.R1040

5'-most EST LIB3049-039-Q1-E1-E1

Seq. No. 9456

Contig ID 16095 1.R1040

5'-most EST LIB3170-029-Q1-K1-G10

Seq. No. 9457

Contig ID 16101_1.R1040 5'-most EST sat701008810.h1

Method BLASTX NCBI GI g2191150



```
BLAST score
                   291
                   4.0e-26
E value
                   91
Match length
                   64
% identity
                   (AF007269) similar to mitochondrial carrier family
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   9458
                   16101 2.R1040
Contig ID
                   txt700734741.hl
5'-most EST
                   9459
Seq. No.
                   16109 1.R1040
Contig ID
                   LIB3049-039-Q1-E1-C4
5'-most EST
                   9460
Seq. No.
                   16110 1.R1040
Contig ID
                   zhf700952664.hl
5'-most EST
                   9461
Seq. No.
                   16110 2.R1040
Contig ID
                   uC-gmflminsoy118e10b1
5'-most EST
                   9462
Seq. No.
                   16111 1.R1040
Contig ID
                   vzy70\overline{0}753767.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4006909
BLAST score
                   346
                   2.0e-32
E value
                   155
Match length
                   52
% identity
NCBI Description (Z99708) putative protein [Arabidopsis thaliana]
Seq. No.
                   9463
                   16111 2.R1040
Contig ID
                   LIB3049-039-Q1-E1-C6
5'-most EST
Seq. No.
                   9464
                   16112 1.R1040
Contig ID
                   LIB3049-039-Q1-E1-C7
5'-most EST
Seq. No.
                   9465
Contig ID
                   16115 1.R1040
                   LIB3049-039-Q1-E1-D1
5'-most EST
```

Method

BLASTX

NCBI GI

g1078074

BLAST score

165

E value

3.0e-11

Match length

106

% identity

42

NCBI Description

probable membrane protein YLR328w - yeast (Saccharomyces

cerevisiae) >gi_662140 (U20618) Ylr328wp [Saccharomyces

cerevisiae]

Seq. No.

9466

Contig ID

16118 1.R1040



```
ncj700980834.h1
5'-most EST
                  BLASTX
Method
                  q2978452
NCBI GI
                  333
BLAST score
                  2.0e-30
E value
                  89
Match length
% identity
                 (AE001274) MCAK; L549.3 [Leishmania major]
NCBI Description
                  9467
Seq. No.
                  16122 1.R1040
Contig ID
                  leu701154824.h1
5'-most EST
                  9468
Seq. No.
                  16122 2.R1040
Contig ID
                  bth700849722.h1
5'-most EST
                  BLASTX
Method
                  g3877951
NCBI GI
                   241
BLAST score
                   3.0e-20
E value
                  133
Match length
                   29
% identity
NCBI Description (Z81555) predicted using Genefinder [Caenorhabditis
                   elegans]
                   9469
Seq. No.
                   16124 1.R1040
Contig ID
                   wrg700790632.h2
5'-most EST
                   BLASTX
Method
                   g3201477
NCBI GI
                   695
BLAST score
                   6.0e-73
E value
                   280
Match length
% identity
                   (AJ006021) putative PRL1 associated protein [Arabidopsis
NCBI Description
                   thaliana]
                   9470
Seq. No.
                   16125 1.R1040
Contig ID
                   LIB3049-039-Q1-E1-C3
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4218126
BLAST score
                   470
E value
                   6.0e-47
                   165
Match length
 % identity
                   (AL035353) putative protein [Arabidopsis thaliana]
NCBI Description
 Seq. No.
                   9471
                   16128 1.R1040
 Contig ID
                   LIB3049-039-Q1-E1-B2
 5'-most EST
                    9472
 Seq. No.
 Contig ID
                   16131 1.R1040
                   LIB3049-039-Q1-E1-B5
 5'-most EST
```

1680

9473

Seq. No.



16132 1.R1040 Contig ID

5'-most EST LIB3049-039-Q1-E1-B6

Method BLASTX NCBI GI g3377509 BLAST score 434 E value 9.0e-43 Match length 96 % identity 80

NCBI Description (AF056027) auxin transport protein REH1 [Oryza sativa]

9474 Seq. No.

Contig ID 16133 1.R1040

LIB3167-077-P1-K2-G7 5'-most EST

Method BLASTX NCBI GI g2191136 BLAST score 431 E value 3.0e-42Match length 142 % identity 57

NCBI Description (AF007269) Similar to UTP-Glucose Glucosyltransferase;

coded for by A. thaliana cDNA T46230; coded for by A.

thaliana cDNA H76538; coded for by A. thaliana cDNA H76290

[Arabidopsis thaliana]

9475 Seq. No.

Contig ID 16135 1.R1040 5'-most EST ssr700560864.h1

Method BLASTX NCBI GI g3152615 BLAST score 219 E value 1.0e-17 142 Match length % identity 40

NCBI Description (AC004482) hypothetical protein [Arabidopsis thaliana]

>gi 3242727 (AC003040) hypothetical protein [Arabidopsis

thaliana]

9476 Seq. No.

Contig ID 16139 1.R1040

5'-most EST jC-gmle01810053h07a1

BLASTX Method NCBI GI g3122638 BLAST score 420 E value 4.0e-82 Match length 218 % identity 35

PP1/PP2A PHOSPHATASES PLEIOTROPIC REGULATOR PRL1 NCBI Description

>gi_1076381_pir__S49820 PRL1 protein - Arabidopsis thaliana >gi_577733_emb_CAA58031_ (X82824) PRL1 [Arabidopsis

thaliana] >gi 577735 emb CAA58032 (X82825) PRL1

[Arabidopsis thaliana] >gi 2244947 emb CAB10369.1 (Z97339)

PRL1 protein - Arabidopsis thaliana

Seq. No. 9477

Contig ID 16144 1.R1040

5'-most EST LIB3170-032-Q1-K1-H9



Contig ID 16144_2.R1040

5'-most EST LIB3072-054-Q1-E1-E9

Seq. No. 9479

Contig ID 16145_1.R1040

5'-most EST uC-gmrominsoy157e03b1

Method BLASTX
NCBI GI g2665890
BLAST score 2241
E value 0.0e+00
Match length 534

% identity 81

NCBI Description (AF035944) calcium-dependent protein kinase [Fragaria x

ananassa]

Seq. No. 9480

Contig ID 16147_1.R1040

5'-most EST LIB3049-034-Q1-E1-F8

Method BLASTX
NCBI GI 9729538
BLAST score 243
E value 3.0e-20
Match length 103
% identity 54

NCBI Description FERREDOXIN-THIOREDOXIN REDUCTASE, VARIABLE CHAIN PRECURSOR

(FTR-V) (FERREDOXIN-THIOREDOXIN REDUCTASE SUBUNIT A) (FTR-A) >gi_2144358_pir__RDSPTA ferredoxin--thioredoxin reductase (EC 1.18.-.-) chain A precursor - spinach >gi_474766_emb_CAA55480_ (X78880) ferredoxin:thioredoxin reductase [Spinacia oleracea] >gi_861138_emb_CAA54408_ (X77163) ferredoxin-thioredoxin reductase SU A [Spinacia

oleracea]

Seq. No. 9481

Contig ID 16147_2.R1040 5'-most EST smc700745464.h1

Method BLASTX
NCBI GI g1362192
BLAST score 195

E value 1.0e-14 Match length 64 % identity 53

NCBI Description ferredoxin-thioredoxin reductase (EC 1.18.-.-) variable

chain - maize

Seq. No. 9482

Contig ID 16148_1.R1040

5'-most EST LIB3170-032-Q1-K1-B5

Seq. No. 9483

Contig ID 16148_2.R1040 5'-most EST vwf700675339.h1

Seq. No. 9484

Contig ID 16148_3.R1040 5'-most EST txt700733039.h1



16157 1.R1040 Contig ID

LIB3055-006-Q1-N1-B11 5'-most EST

9486 Seq. No.

Contig ID 16158 1.R1040

5'-most EST LIB3049-038-Q1-E1-G1

9487 Seq. No.

16162 1.R1040 Contig ID

LIB3049-038-Q1-E1-G2 5'-most EST

BLASTX Method g1421730 NCBI GI BLAST score 284 2.0e-25 E value 74 Match length 77 % identity

NCBI Description (U43082) RF2 [Zea mays]

9488 Seq. No.

16163 1.R1040 Contig ID pmv700891217.h1 5'-most EST

Method BLASTN q2281081 NCBI GI BLAST score 46 1.0e-16 E value Match length 46 100 % identity

Arabidopsis thaliana chromosome II BAC F18019 genomic NCBI Description

sequence, complete sequence [Arabidopsis thaliana]

9489 Seq. No.

16163 3.R1040 Contig ID

LIB3072-023-Q1-E1-D8 5'-most EST

Method BLASTN g2281081 NCBI GI BLAST score 46 9.0e-17 E value 46 Match length 100 % identity

Arabidopsis thaliana chromosome II BAC F18019 genomic NCBI Description

sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 9490

16165 1.R1040 Contig ID awf700839728.h1 5'-most EST

BLASTX Method g3287683 NCBI GI 778 BLAST score 1.0e-82 E value 337 Match length 15 % identity

(AC003979) Similar to apoptosis protein MA-3 gb_D50465 from NCBI Description

Mus musculus. [Arabidopsis thaliana]

9491 Seq. No.



Contig ID 16170_1.R1040

5'-most EST LIB3170-030-Q1-K1-H9

Method BLASTX
NCBI GI g2583123
BLAST score 450
E value 6.0e-45
Match length 122
% identity 75

NCBI Description (AC002387) putative nucleotide sugar epimerase [Arabidopsis

thaliana]

Seq. No. 9492

Contig ID 16171 1.R1040

5'-most EST LIB3170-032-Q1-K1-A5

Method BLASTX
NCBI GI 9730692
BLAST score 234
E value 2.0e-19
Match length 122
% identity 44

NCBI Description RUBISCO-ASSOCIATED PROTEIN >gi_454179 (L28804) putative

[Glycine max] >gi_1090839_prf__2019481A RuBisCO complex

protein [Glycine max]

Seq. No. 9493

Contig ID 16173 1.R1040

5'-most EST LIB3109-028-Q1-K1-D6

Method BLASTX
NCBI GI 94325282
BLAST score 362
E value 2.0e-47
Match length 109
% identity 86

NCBI Description (AF123310) NAC domain protein NAM [Arabidopsis thaliana]

>gi 4325286 gb AAD17314 (AF123311) NAC domain protein NAM

[Arabidopsis thaliana]

Seq. No. 9494

Contig ID 16176_1.R1040

5'-most EST LIB3170-030-Q1-K1-A9

Seq. No. 9495

Contig ID 16177 1.R1040

5'-most EST jC-gmst02400065b11a2

Seq. No. 9496

Contig ID 16177_2.R1040 5'-most EST zhf700952970.h1

Seq. No. 9497

Contig ID 16180 1.R1040

5'-most EST LIB3049-038-Q1-E1-E8

Seq. No. 9498

Contig ID 16183 1.R1040

5'-most EST LIB3106-092-Q1-K1-F6



```
Seq. No.
                   9499
                  16186 1.R1040
Contig ID
5'-most EST
                  ssr700553976.hl
                  BLASTX
Method
NCBI GI
                  g1171577
BLAST score
                   337
                  2.0e-31
E value
                  166
Match length
                   42
% identity
NCBI Description
                 (X95343) hypersensitivity-related gene [Nicotiana tabacum]
                  9500
```

16188 1.R1040 Contig ID

5'-most EST LIB3170-031-Q1-J1-D9

9501 Seq. No.

16198 1.R1040 Contig ID

5'-most EST LIB3049-038-Q1-E1-D1

Method BLASTN NCBI GI g558628 BLAST score 441 E value 0.0e + 00978 Match length 90 % identity

NCBI Description P.vulgaris mRNA for RNP1 chloroplast RNA binding protein

Seq. No. 9502

16198 2.R1040 Contig ID

LIB3106-108-Q1-K1-A2 5'-most EST

Method BLASTN NCBI GI q558628 BLAST score 71 E value 1.0e-31

Match length 285 87 % identity

NCBI Description P.vulgaris mRNA for RNP1 chloroplast RNA binding protein

Seq. No. 9503

16200 1.R1040 Contig ID

5'-most EST LIB3170-032-Q1-K1-A4

Seq. No. 9504

16200 2.R1040 Contig ID 5'-most EST leu701157014.h1

Seq. No. 9505

16201 1.R1040 Contig ID

5'-most EST LIB3170-077-Q1-K1-F12

Method BLASTX NCBI GI g404690 BLAST score 1290 E value 1.0e-142 Match length 484 % identity 51

(L19075) cytochrome P450 [Catharanthus roseus] NCBI Description

Seq. No. 9506



```
16201 3.R1040
Contig ID
                   seb700653537.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4376203
                   210
BLAST score
                   9.0e-17
E value
                   77
Match length
                   47
% identity
                   (U35226) putative cytochrome P-450 [Nicotiana
NCBI Description
                   plumbaginifolia]
                   9507
Seq. No.
                   16204 1.R1040
Contig ID
                   nci700981150.hl
5'-most EST
                   BLASTX
Method
                   q3687235
NCBI GI
                   560
BLAST score
                   3.0e-57
E value
                   149
Match length
                   75
% identity
                   (AC005169) putative copia-like transposable element
NCBI Description
                   [Arabidopsis thaliana]
                   9508
Seq. No.
                   16204 2.R1040
Contig ID
                   uC-gmrominsoy206f05b1
5'-most EST
                   BLASTX
Method
                   a3687235
NCBI GI
                   382
BLAST score
                   1.0e-36
E value
Match length
                   116
                   68
% identity
                   (AC005169) putative copia-like transposable element
NCBI Description
                    [Arabidopsis thaliana]
                   9509
Seq. No.
                   16204 3.R1040
Contig ID
                   crh700854648.h1
5'-most EST
Method
                   BLASTX
                   q3687235
NCBI GI
BLAST score
                   145
                   3.0e-09
E value
                    44
Match length
                    64
 % identity
                    (AC005169) putative copia-like transposable element
NCBI Description
                    [Arabidopsis thaliana]
                    9510
 Seq. No.
                    16208 1.R1040
 Contig ID
                   LIB3170-078-Q1-K1-D8
 5'-most EST
                    BLASTX
 Method
                    g1483150
 NCBI GI
                    450
 BLAST score
                    1.0e-44
 E value
                    109
 Match length
                    78
 % identity
                   (D84417) monodehydroascorbate reductase [Arabidopsis
 NCBI Description
```

% identity

NCBI Description



thaliana]

```
Seq. No.
                   9511
                   16208 2.R1040
Contig ID
                   LIB3109-046-Q1-K1-A11
5'-most EST
                   BLASTX
Method
                   g1483150
NCBI GI
BLAST score
                   263
E value
                   7.0e-23
Match length
                   54
% identity
                   (D84417) monodehydroascorbate reductase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   9512
                   16209 1.R1040
Contig ID
                   k1170\overline{1}214815.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4335734
BLAST score
                   386
                   7.0e-37
E value
                   182
Match length
% identity
                   45
                   (AC006248) putative calmodulin [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   16213 1.R1040
Contig ID
                   zhf700959796.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4262240
BLAST score
                   535
                   2.0e-54
E value
                   139
Match length
% identity
                   (AC006200) putative stress protein [Arabidopsis thaliana]
NCBI Description
                    9514
Seq. No.
                    16213 2.R1040
Contig ID
5'-most EST
                    k1170\overline{1}214913.h1
                    BLASTX
Method
                    q4262240
NCBI GI
                    487
BLAST score
                    4.0e-49
E value
                    161
Match length
% identity
                   (AC006200) putative stress protein [Arabidopsis thaliana]
NCBI Description
                    9515
Seq. No.
                    16228 1.R1040
Contig ID
                    sat701012934.h1
5'-most EST
                    BLASTX
Method
NCBI GI
                    g2190550
BLAST score
                    1063
                    1.0e-116
E value
Match length
                    274
```

(ACO01229) ESTs gb_T45673,gb_N37512 come from this gene.



[Arabidopsis thaliana]

```
9516
Seq. No.
                  16230 1.R1040
Contig ID
                  rry700808402.hl
5'-most EST
                  BLASTX
Method
                  g4115934
NCBI GI
BLAST score
                  614
E value
                  1.0e-63
                  137
Match length
                  77
% identity
                   (AF118223) contains similarity to Methanobacterium
NCBI Description
                  thermoautotrophicum transcriptional regulator (GB:AE000850)
                   [Arabidopsis thaliana]
                   9517
Seq. No.
                   16247 1.R1040
Contig ID
                   trc700561396.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                   q3075402
                   304
BLAST score
                   4.0e-27
E value
Match length
                   104
% identity
                  (AC004484) unknown protein [Arabidopsis thaliana]
NCBI Description
                   9518
Seq. No.
                   16247 2.R1040
Contig ID
                   smc700747392.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3075402
BLAST score
                   207
E value
                   4.0e-16
Match length
                   79
                   53
% identity
                   (AC004484) unknown protein [Arabidopsis thaliana]
NCBI Description
                   9519
Seq. No.
                   16248 1.R1040
Contig ID
                   LIB3050-010-Q1-E1-F2
5'-most EST
Method
                   BLASTX
                   g1364059
NCBI GI
                   307
BLAST score
                   6.0e-28
E value
Match length
                   129
% identity
                   46
                   dioscorin class A precursor - Dioscorea cayenensis
NCBI Description
                   >gi_433463_emb_CAA53781_ (X76187) storage protein
                   [Dioscorea cayenensis]
                   9520
Seq. No.
                   16252 1.R1040
Contig ID
5'-most EST
                   uC-gmropic045h07b1
```

Seq. No. 9521

Contig ID 16253_1.R1040

5'-most EST LIB3049-004-Q1-E1-H3



Contig ID 16264_1.R1040

5'-most EST LIB3106-093-Q1-K1-G8

Method BLASTX
NCBI GI g2880049
BLAST score 235
E value 1.0e-19
Match length 91
% identity 55

NCBI Description (AC002340) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9523

Contig ID 16268 1.R1040

5'-most EST LIB3170-067-Q1-K1-G10

85

Method BLASTX
NCBI GI g1174867
BLAST score 334
E value 5.0e-31
Match length 72

% identity

NCBI Description UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING

PROTEIN QP-C (UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 8.2

KD PROTEIN) >gi_633687_emb_CAA55862 (X79275)

ubiquinol--cytochrome c reductase [Solanum tuberosum]

>gi_1094912_prf__2107179A cytochrome c
oxidase:SUBUNIT=8.2kD [Solanum tuberosum]

Seq. No. 9524

Contig ID 16268_2.R1040 5'-most EST hrw701060207.h1

Method BLASTX
NCBI GI g1174867
BLAST score 267
E value 2.0e-23
Match length 67
% identity 78

NCBI Description UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING

PROTEIN QP-C (UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 8.2

KD PROTEIN) >gi_633687_emb_CAA55862 (X79275)

ubiquinol--cytochrome c reductase [Solanum tuberosum]

>gi_1094912_prf__2107179A cytochrome c oxidase:SUBUNIT=8.2kD [Solanum tuberosum]

Seq. No. 9525

Contig ID 16270 1.R1040

5'-most EST LIB3106-069-P1-K1-G3

Method BLASTX
NCBI GI g2104536
BLAST score 318
E value 3.0e-29
Match length 120
% identity 53

NCBI Description (AF001308) predicted glycosyl transferase [Arabidopsis

thaliana]

Seq. No. 9526



```
16272 1.R1040
Contig ID
                  LIB3170-028-Q1-K1-E3
5'-most EST
Method
                  BLASTX
                   g1684851
NCBI GI
                   265
BLAST score
                   5.0e-24
E value
                   91
Match length
% identity
                   64
NCBI Description (U77935) DnaJ-like protein [Phaseolus vulgaris]
Seq. No.
                   9527
                   16288 1.R1040
Contig ID
                   LIB3170-030-Q1-K1-G6
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3123349
BLAST score
                   295
E value
                   1.0e-26
                   74
Match length
% identity
                   77
                   (AJ005788) hypothetical protein [Cicer arietinum]
NCBI Description
Seq. No.
                   16293 1.R1040
Contig ID
                   LIB3170-029-Q1-K1-B2
5'-most EST
                   BLASTN
Method
NCBI GI
                   g4220468
BLAST score
                   45
                   5.0e-16
E value
Match length
                   113
% identity
                   85
                   Arabidopsis thaliana chromosome II BAC T8011 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
Seq. No.
                   9529
                   16301 1.R1040
Contig ID
                   LIB3049-035-Q1-E1-H6
5'-most EST
                   BLASTX
Method
                   g4335745
NCBI GI
BLAST score
                   200
                   2.0e-15
E value
Match length
                   86
                   50
% identity
                   (AC006284) putative hydrolase (contains an
NCBI Description
                   esterase/lipase/thioesterase active site serine domain
                   (prosite: PS50187) [Arabidopsis thaliana]
Seq. No.
                   9530
                   16304 1.R1040
Contig ID
                   uC-gmrominsoy071d08b1
5'-most EST
                   9531
Seq. No.
                   16304 2.R1040
Contig ID
                   jC-gmf102220057a02a1
5'-most EST
Seq. No.
                   9532
```

16312 1.R1040

leu701157523.h1

Contig ID

5'-most EST



Method BLASTX NCBI GI q170131 352 BLAST score E value 3.0e-33 Match length 88 76 % identity

NCBI Description (M55322) ribosomal protein 30S subunit [Spinacia oleracea]

Seq. No.

9533

Contig ID

16317 1.R1040

5'-most EST

LIB3170-029-Q1-K1-C1

Seq. No.

9534

Contig ID

16321 1.R1040

5'-most EST

LIB3049-035-Q1-E1-F9

Seq. No.

9535

Contig ID

16324 1.R1040

5'-most EST

LIB3073-021-Q1-K1-F12

Method NCBI GI BLAST score BLASTX g2190548 244

E value Match length 2.0e-20 68

% identity NCBI Description 66

(AC001229) EST gb_ATTS1121 comes from this gene.

[Arabidopsis thaliana]

Seq. No.

9536

Contig ID

16324 2.R1040

5'-most EST

LIB3073-021-Q1-K1-F11

Method BLASTX NCBI GI g2190548 BLAST score 359 E value 5.0e-34 Match length 106

% identity NCBI Description

(AC001229) EST gb ATTS1121 comes from this gene.

[Arabidopsis thaliana]

Seq. No.

9537

62

Contig ID

16328 1.R1040

5'-most EST

jC-gmst02400007d01a1

Seq. No.

9538

Contig ID 5'-most EST

16335 2.R1040 awf700841716.hl

Seq. No.

9539

Contig ID

16337 1.R1040

5'-most EST

 $LIB31\overline{7}0-031-Q1-K1-A5$

Method BLASTX NCBI GI g1588365 BLAST score 185 E value 1.0e-13

Match length 74 % identity 46



```
NCBI Description signal peptidase:SUBUNIT=12kD [Homo sapiens]
                  9540
Seq. No.
                  16345 1.R1040
Contig ID
                  crh700856185.h1
5'-most EST
                   9541
Seq. No.
                   16346 1.R1040
Contig ID
5'-most EST
                  LIB3049-035-Q1-E1-E12
                   9542
Seq. No.
                   16355 1.R1040
Contig ID
                   leu701148285.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4510363
                   623
BLAST score
                   1.0e-64
E value
                   134
Match length
% identity
                   (AC007017) putative DNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   9543
                   16355 2.R1040
Contig ID
                   uC-gmropic022c11b1
5'-most EST
Method
                   BLASTX
                   q4510363
NCBI GI
BLAST score
                   503
                   8.0e-51
E value
Match length
                   106
                   89
% identity
                   (AC007017) putative DNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   9544
                   16355 3.R1040
Contig ID
                   trc700565054.hl
5'-most EST
                   BLASTX
Method
                   g4510363
NCBI GI
                   700
BLAST score
                   9.0e-74
E value
                   146
Match length
% identity
                   (AC007017) putative DNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
                   9545
Seq. No.
                   16355 4.R1040
Contig ID
5'-most EST
                   trc700565456.h1
                   BLASTN
Method
NCBI GI
                   q4510360
```

NCBI GI g4510360 BLAST score 49 E value 2.0e-18 Match length 105 % identity 87

NCBI Description Arabidopsis thaliana chromosome II BAC F11F19 genomic sequence, complete sequence



```
9546
Seq. No.
                  16355 5.R1040
Contig ID
                  LIB3074-030-Q1-K2-B8
5'-most EST
                  BLASTX
Method
                  g4510363
NCBI GI
                  427
BLAST score
E value
                   4.0e-42
Match length
                  93
                  80
% identity
                   (AC007017) putative DNA-binding protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   9547
                   16355 6.R1040
Contig ID
                   epx701110325.hl
5'-most EST
Method
                   BLASTN
                   g4510360
NCBI GI
BLAST score
                   65
E value
                   3.0e-28
Match length
                   109
                   90
% identity
                   Arabidopsis thaliana chromosome II BAC F11F19 genomic
NCBI Description
                   sequence, complete sequence
Seq. No.
                   9548
                   16359 1.R1040
Contig ID
                   LIB3170-029-Q1-K1-A2
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3386596
BLAST score
                   163
                   1.0e-11
E value
                   47
Match length
% identity
                   68
                   (AC004665) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi_3702346 (AC005397) unknown protein [Arabidopsis
                   thaliana]
                   9549
Seq. No.
                   16370 1.R1040
Contig ID
                   LIB3170-030-Q1-K1-F4
5'-most EST
                   BLASTX
Method
                   g2244818
NCBI GI
BLAST score
                   356
E value
                   6.0e-34
                   106
Match length
                   73
% identity
                   (Z97336) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   9550
Seq. No.
                   16374 1.R1040
Contig ID
                   LIB3170-030-Q1-K1-E4
5'-most EST
```

9551

16379_1.R1040 vzy700753123.h1

Seq. No. Contig ID

5'-most EST



9552 Seq. No. 16379 2.R1040 Contig ID LIB3107-055-Q1-K1-F8 5'-most EST 9553 Seq. No. 16387 1.R1040 Contig ID pxt700944294.hl 5'-most EST Method BLASTX g629670 NCBI GI 223 BLAST score 3.0e-18 E value 84 Match length 58 % identity NCBI Description hypothetical protein - tomato 9554 Seq. No. 16391 1.R1040 Contig ID uC-gmrominsoy060b04b1 5'-most EST Method BLASTX g1935019 NCBI GI 526 BLAST score 2.0e-53 E value 190 Match length 59 % identity NCBI Description (Z93774) sucrose transport protein [Vicia faba] 9555 Seq. No. 16393 1.R1040 Contig ID crh700854076.hl 5'-most EST BLASTX Method g2673917 NCBI GI BLAST score 609 2.0e-79 E value Match length 206 35 % identity (AC002561) putative ATP-dependent RNA helicase [Arabidopsis NCBI Description thaliana] 9556 Seq. No. 16397 1.R1040 Contig ID LIB3050-008-Q1-E1-C1 5'-most EST Method BLASTX NCBI GI q2765240 BLAST score 193 9.0e-15 E value 80 Match length % identity (Y12805) invertase inhibitor [Nicotiana tabacum] NCBI Description Seq. No. 9557 16398 1.R1040 Contig ID uC-gmropic055e11b1 5'-most EST Method BLASTX NCBI GI g4335750

1694

256

90

2.0e-21

BLAST score

Match length

E value



% identity 50

NCBI Description (AC006284) putative beta-1,3-endoglucanase [Arabidopsis

thaliana]

Seq. No. 9558

Contig ID 16398_2.R1040

5'-most EST LIB3050-008-Q1-E1-B3

Method BLASTX
NCBI GI g4335750
BLAST score 219
E value 1.0e-17
Match length 64

% identity 56 NCBI Description (AC006284) putative beta-1,3-endoglucanase [Arabidopsis

thaliana]

Seq. No. 9559

Contig ID 16403 1.R1040

5'-most EST LIB3072-049-Q1-K1-E7

Seq. No. 9560

Contig ID 16406_1.R1040 5'-most EST uC-gmropic106g10b1

Seq. No. 9561

Contig ID 16407 1.R1040

5'-most EST LIB3109-007-Q1-K1-D8

Method BLASTX
NCBI GI g1170089
BLAST score 637
E value 3.0e-66
Match length 212
% identity 57

NCBI Description

GLUTATHIONE S-TRANSFERASE ERD13 (CLASS PHI)

>gi 481822_pir__S39542 probable glutathione transferase (EC

2.5.1.18) (clone ERD13) - Arabidopsis thaliana

>gi_497789_dbj_BAA04554_ (D17673) glutathione S-transferase
[Arabidopsis thaliana] >gi_3201614 (AC004669) glutathione

S-transferase [Arabidopsis thaliana]

Seq. No. 9562

Contig ID 16407_2.R1040

5'-most EST g4302960
Method BLASTX
NCBI GI g3201613
BLAST score 361
E value 5.0e-34
Match length 130
% identity 55

NCBI Description (AC004669) glutathione S-transferase [Arabidopsis thaliana]

Seq. No. 9563

Contig ID 16410 1.R1040 5'-most EST seb700650214.h1

Seq. No. 9564

Contig ID 16410_2.R1040

% identity

40



```
LIB3170-029-Q1-K1-G6
5'-most EST
                  9565
Seq. No.
                  16414 1.R1040
Contig ID
                   jex70\overline{0}908176.h1
5'-most EST
                  BLASTX
Method
                  g3738297
NCBI GI
                  201
BLAST score
                  1.0e-15
E value
                   67
Match length
                   24
% identity
NCBI Description (AC005309) unknown protein [Arabidopsis thaliana]
                   9566
Seq. No.
                   16415 1.R1040
Contiq ID
                   LIB3049-034-Q1-E1-E5
5'-most EST
                   9567
Seq. No.
                   16422 1.R1040
Contig ID
                   sat701007849.h1
5'-most EST
                   BLASTX
Method
                   g3955021
NCBI GI
                   304
BLAST score
                   1.0e-27
E value
                   72
Match length
                   78
% identity
                   (AJ010811) HB2 homeodomain protein [Populus tremula x
NCBI Description
                   Populus tremuloides]
                   9568
Seq. No.
                   16425 1.R1040
Contig ID
                   LIB3049-034-Q1-E1-C10
5'-most EST
                   BLASTX
Method
                   q3080398
NCBI GI
                   246
BLAST score
                   6.0e-21
E value
                   112
Match length
                   24
% identity
                   (AL022603) putative protein [Arabidopsis thaliana]
NCBI Description
                    9569
Seq. No.
                    16426 1.R1040
Contig ID
5'-most EST
                   uC-gmronoir007b08b1
                    9570
Seq. No.
                    16426 2.R1040
 Contig ID
                    asn701138153.hl
 5'-most EST
 Seq. No.
                    9571
                    16428 1.R1040
 Contig ID
                    awf700836372.h1
 5'-most EST
                    BLASTX
 Method
                    g2959370
 NCBI GI
 BLAST score
                    318
 E value
                    8.0e-29
 Match length
                    139
```



NCBI Description (AL022117) hypothetical protein [Schizosaccharomyces pombe]

Seq. No. 9572

Contig ID 16428 2.R1040 5'-most EST dpv701101125.h1

Method BLASTN
NCBI GI g3510343
BLAST score 49
E value 2.0e-18
Match length 89
% identity 89

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MJC20, complete sequence [Arabidopsis thaliana]

Seq. No. 9573

Contig ID 16435 1.R1040

5'-most EST LIB3092-043-Q1-K1-C12

Method BLASTX
NCBI GI g2673920
BLAST score 193
E value 2.0e-14
Match length 87
% identity 48

NCBI Description (AC002561) similar to Drosophila couch potato protein

[Arabidopsis thaliana]

Seq. No. 9574

Contig ID 16440_1.R1040 5'-most EST zzp700832734.h1

Seq. No. 9575

Contig ID 16445_1.R1040 5'-most EST sat701011620.h1

Method BLASTX
NCBI GI g132944
BLAST score 469
E value 4.0e-47
Match length 92
% identity 92

NCBI Description 60S RIBOSOMAL PROTEIN L3 >gi_81658_pir__JQ0772 ribosomal

protein L3 (ARP2) - Arabidopsis thaliana >gi_806279 (M32655) ribosomal protein [Arabidopsis thaliana]

Seq. No. 9576

Contig ID 16447 1.R1040 5'-most EST zhf700952425.h1

Method BLASTX
NCBI GI g2104691
BLAST score 291
E value 1.0e-25
Match length 139
% identity 41

NCBI Description (U92794) alpha glucosidase II, beta subunit [Mus musculus]

Seq. No. 9577

Contig ID 16453_1.R1040 5'-most EST fC-gmse7000757563f1



Contig ID 16453_2.R1040 5'-most EST zhf700963836.h1

Seq. No. 9579

Contig ID 16453_3.R1040 5'-most EST zsg701127106.h1

Seq. No. 9580

Contig ID 16453_4.R1040

5'-most EST jC-gmro02910064h10d1

Seq. No. 9581

Contig ID 16453_5.R1040 5'-most EST fC-gmro700848077d4

Seq. No. 9582

Contig ID 16453_7.R1040 5'-most EST leu701144767.h1

Seq. No. 9583

Contig ID 16466 1.R1040 5'-most EST sat701006587.h1

Method BLASTX
NCBI GI g729704
BLAST score 250
E value 4.0e-21
Match length 135
% identity 8

NCBI Description DNA-BINDING PROTEIN HEXBP (HEXAMER-BINDING PROTEIN)

>gi_1078700_pir__A47156 hexamer-binding protein HEXBP -Leishmania major >gi 159342 (M94390) HEXBP DNA binding

protein [Leishmania major]

Seq. No. 9584

Contig ID 16469 1.R1040 5'-most EST kl1701206014.h1

Method BLASTX
NCBI GI g3292827
BLAST score 746
E value 4.0e-79
Match length 158
% identity 88

NCBI Description (AL031018) putative protein [Arabidopsis thaliana]

Seq. No. 9585

Contig ID 16473 1.R1040

5'-most EST LIB3170-030-Q1-K1-H3

Seq. No. 9586

Contig ID 16486_1.R1040 5'-most EST gsv701049029.h1

Method BLASTX
NCBI GI g2911081
BLAST score 149
E value 5.0e-09



Match length 257 % identity 24

NCBI Description (AL021960) putative protein [Arabidopsis thaliana]

Seq. No.

Contig ID 16487_1.R1040

5'-most EST LIB3170-029-Q1-K1-A6

9587

Seq. No. 9588

Contig ID 16492 1.R1040

5'-most EST LIB3049-033-Q1-E1-D6

Method BLASTX
NCBI GI g2621299
BLAST score 191
E value 3.0e-14
Match length 140
% identity 34

NCBI Description (AE000812) conserved protein [Methanobacterium

thermoautotrophicum]

Seq. No. 9589

Contig ID 16499 1.R1040 5'-most EST kl1701209813.h1

Seq. No. 9590

Contig ID 16500_1.R1040

5'-most EST LIB3049-033-Q1-E1-E3

Seq. No. 9591

Contig ID 16505_1.R1040

5'-most EST jC-gmro02910037f04d1

Method BLASTX
NCBI GI g1706318
BLAST score 2055
E value 0.0e+00
Match length 501
% identity 77

NCBI Description GLUTAMATE DECARBOXYLASE (GAD) (ERT D1)

>gi_1362098_pir__ S56177 probable glutamate decarboxylase tomato >gi_995555_emb_CAA56812_ (X80840) homology to
pyroxidal-5'-phosphate-dependent glutamate decarboxylases;

putative start codon [Lycopersicon esculentum]

Seq. No. 9592

Contig ID 16506 1.R1040

5'-most EST LIB3049-033-Q1-E1-E8

Method BLASTX
NCBI GI 94432846
BLAST score 171
E value 4.0e-12
Match length 105
% identity 49

NCBI Description (AC006283) unknown protein [Arabidopsis thaliana]

Seq. No. 9593

Contig ID 16508 1.R1040 5'-most EST leu701146421.h1

1699



Contig ID 16509_2.R1040

5'-most EST LIB3049-028-Q1-E1-C8

Method BLASTX
NCBI GI g3080439
BLAST score 356
E value 7.0e-34
Match length 96
% identity 73

NCBI Description (AL022605) putative protein [Arabidopsis thaliana]

Seq. No. 9595

Contig ID 16513_1.R1040 5'-most EST gsv701056187.h1

Seq. No. 9596

Contig ID 16519 1.R1040

5'-most EST LIB3049-033-Q1-E1-C7

Seq. No. 9597

Contig ID 16525_1.R1040

5'-most EST LIB3170-029-Q1-K1-F5

Seq. No. 9598

Contig ID 16527_1.R1040 5'-most EST pxt700941561.h1

Seq. No. 9599

Contig ID 16531_2.R1040 5'-most EST jsh701064262.h1

Seq. No. 9600

Contig ID 16531_3.R1040

5'-most EST LIB3065-014-Q1-N1-A7

Seq. No. 9601

Contig ID 16531 4.R1040

5'-most EST LIB3049-003-Q1-E1-A10

Method BLASTN
NCBI GI g498704
BLAST score 33
E value 7.0e-09
Match length 53
% identity 91

NCBI Description S.oleracea slfa mRNA

Seq. No. 9602

Contig ID 16532_1.R1040 5'-most EST kl1701206566.h1

Seq. No. 9603

Contig ID 16533_1.R1040

5'-most EST LIB3049-033-Q1-E1-A6

Seq. No. 9604

Contig ID 16547_1.R1040

1700



5'-most EST asn701140281.h1

Seq. No. 9605

Contig ID 16548_1.R1040

5'-most EST LIB3109-055-Q1-K1-C8

Method BLASTX
NCBI GI g2160182
BLAST score 166
E value 4.0e-11
Match length 81
% identity 43

NCBI Description (AC000132) ESTs gb ATTS1236, gb T43334, gb N97019, gb AA395203

come from this gene. [Arabidopsis thaliana]

Seq. No. 9606

Contig ID 16549 1.R1040 5'-most EST g5175367 Method BLASTX NCBI GI g124429 BLAST score 1339 E value 1.0e-148 Match length 342 % identity 75

NCBI Description 37 KD CHLOROPLAST INNER ENVELOPE MEMBRANE PROTEIN PRECURSOR

(E37) >gi_99543_pir__S14409 membrane protein, 37K, precursor, chloroplast inner envelope - spinach

>gi 21228 emb CAA40283 (X56963) 37 kD inner envelope

membrane polypeptide [Spinacia oleracea]

Seq. No. 9607

Contig ID 16549_3.R1040 5'-most EST kmv700740656.h1

Seq. No. 9608

Contig ID 16559_1.R1040 5'-most EST zhf700954620.h1

Seq. No. 9609

Contig ID 16563 1.R1040

5'-most EST LIB3109-033-Q1-K1-A1

Method BLASTN
NCBI GI g3819163
BLAST score 551
E value 0.0e+00
Match length 591
% identity 98

NCBI Description Glycine max cctd gene

Seq. No. 9610

Contig ID 16571_1.R1040

5'-most EST LIB3049-002-Q1-E1-G9

Method BLASTX
NCBI GI g3599491
BLAST score 1465
E value 1.0e-163
Match length 338
% identity 80

NCBI GI

BLAST score



```
NCBI Description
                  (AF085149) putative aminotransferase [Capsicum chinense]
Seq. No.
                  9611
                  16573 1.R1040
Contig ID
5'-most EST
                  LIB3049-032-Q1-E1-G10
Method
                  BLASTX
NCBI GI
                  g4538981
BLAST score
                  162
E value
                  6.0e-11
Match length
                  59
                  56
% identity
                  (AL049487) putative protein [Arabidopsis thaliana]
NCBI Description
                  9612
Seq. No.
                  16574 1.R1040
Contig ID
5'-most EST
                  leu701149875.h1
Seq. No.
                  9613
                  16574 2.R1040
Contig ID
                  LIB3107-077-Q1-K1-E3
5'-most EST
Seq. No.
                  16576 1.R1040
Contig ID
                  LIB3049-032-Q1-E1-G12
5'-most EST
                  9615
Seq. No.
Contig ID
                  16576 2.R1040
                  xpa700795012.h1
5'-most EST
Seq. No.
                  9616
                  16581 1.R1040
Contig ID
5'-most EST
                  LIB3074-040-Q1-K1-C2
Method
                  BLASTN
NCBI GI
                  q1666172
BLAST score
                  101
E value
                  2.0e-49
Match length
                  289
% identity
                  84
NCBI Description N.plumbaginifolia mRNA for BTF3-like transcription factor
Seq. No.
                   9617
Contig ID
                   16582 1.R1040
5'-most EST
                  rlr700897329.h1
                  BLASTX
Method
NCBI GI
                  g3094014
BLAST score
                  160
                   4.0e-17
E value
Match length
                   91
% identity
                   51
NCBI Description
                  (AF060862) unknown [Homo sapiens]
Seq. No.
                   9618
                  16582 3.R1040
Contig ID
5'-most EST
                  LIB3049-032-Q1-E1-D11
                  BLASTX
Method
```

g3094014 206



E value 3.0e-16 Match length 71 % identity 54

NCBI Description

(AF060862) unknown [Homo sapiens]

Seq. No.

9619

Contig ID

16593_1.R1040

5'-most EST

LIB3170-029-Q1-K1-B5

Seq. No.

9620

Contig ID

16610 1.R1040

5'-most EST

LIB3051-058-Q1-K2-G8

Method NCBI GI BLASTX g2833469 143

BLAST score E value Match length

1.0e-08 119

32

% identity NCBI Description

HYPOTHETICAL 22.2 KD PROTEIN SLR0305

>gi_1001792_dbj_BAA10672_ (D64005) hypothetical protein

[Synechocystis sp.]

Seq. No. Contig ID

No. 9621

16631 1.R1040

5'-most EST

LIB3139-094-P1-N1-E8

Method NCBI GI BLASTX g4559384

BLAST score E value 924 1.0e-100

Match length

244

% identity NCBI Description

73 (AC006526) unknown protein [Arabidopsis thaliana]

Seq. No.

9622

Contig ID

16631_2.R1040

5'-most EST

g5509744

Seq. No.

9623

Contig ID 5'-most EST 16631_3.R1040 wrg700787781.h2

Seq. No.

9624

Contig ID 5'-most EST

16637_1.R1040 asn701133352.h1

Seq. No.

9625

Contig ID

16640 1.R1040

5'-most EST

LIB3049-031-Q1-E1-G3

Seq. No.

9626

Contig ID 5'-most EST

16644_1.R1040 wvk700681309.h2

Method NCBI GI BLAST score BLASTX g3510261

397 2.0e-38

E value 2.00 Match length 127

5-30



% identity

(AC005310) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No.

16644 2.R1040 Contig ID

5'-most EST jC-qmro02910016f10a1

BLASTX Method q3522950 NCBI GI 292 BLAST score 2.0e-26 E value Match length 99 62 % identity

(AC004411) hypothetical protein [Arabidopsis thaliana] NCBI Description

9628 Seq. No.

16648 1.R1040 Contig ID

LIB3049-031-Q1-E1-D8 5'-most EST

Seq. No. 9629

16649 1.R1040 Contig ID $k1170\overline{1}213122.h1$ 5'-most EST

Seq. No. 9630

16655 1.R1040 Contig ID bth700847436.hl 5'-most EST

9631 Seq. No.

16656 1.R1040 Contig ID

LIB3049-031-Q1-E1-E6 5'-most EST

BLASTX Method q4558564 NCBI GI 284 BLAST score E value 2.0e-25 Match length 129 48

% identity

(AC007138) predicted protein of unknown function NCBI Description

[Arabidopsis thaliana]

9632 Seq. No.

Contig ID 16657 1.R1040

5'-most EST LIB3093-001-Q1-K1-F1

Seq. No. 9633

Contig ID 16660 1.R1040

LIB3107-030-Q1-K1-H10 5'-most EST

Method BLASTX q3193296 NCBI GI BLAST score 163 5.0e-11 E value Match length 119 % identity

(AF069298) similar to pectinesterase [Arabidopsis thaliana] NCBI Description

Seq. No. 9634

16664 1.R1040 Contig ID

LIB3049-031-Q1-E1-F12 5'-most EST



Contig ID 16678_1.R1040 5'-most EST fC-gmse700667261f4

Method BLASTX
NCBI GI g3297819
BLAST score 787
E value 7.0e-84
Match length 169
% identity 88

NCBI Description (AL031032) protein kinase - like protein [Arabidopsis

thaliana]

Seq. No. 9636

Contig ID 16678 2.R1040

5'-most EST LIB3049-031-Q1-E1-D4

Method BLASTX
NCBI GI g3297819
BLAST score 462
E value 4.0e-46
Match length 105
% identity 84

NCBI Description (AL031032) protein kinase - like protein [Arabidopsis

thaliana]

Seq. No. 9637

Contig ID 16682 1.R1040

5'-most EST LIB3049-031-Q1-E1-A5

Seq. No. 9638

Contig ID 16684_1.R1040

5'-most EST LIB3106-105-Q1-K1-A9

Seq. No. 9639

Contig ID 16684_2.R1040

5'-most EST LIB3107-008-Q1-K1-C12

Seq. No. 9640

Contig ID 16688_1.R1040 5'-most EST ary700764359.h1

Method BLASTX
NCBI GI g3420052
BLAST score 615
E value 4.0e-64
Match length 150
% identity 77

NCBI Description (AC004680) putative ubiqinone reductase [Arabidopsis

thaliana]

Seq. No. 9641

Contig ID 16690 1.R1040

5'-most EST LIB3049-031-Q1-E1-B2

Seq. No. 9642

Contig ID 16691 1.R1040

5'-most EST LIB3049-031-Q1-E1-B3

Method BLASTX NCBI GI g1723832

1705



BLAST score 181 E value 2.0e-13 Match length 93 % identity 38

NCBI Description HYPOTHETICAL 15.9 KD PROTEIN IN OLE1-DUP1 INTERGENIC REGION >gi 2132531 pir S64058 probable membrane protein YGL054c -

yeast (Saccharomyces cerevisiae) >gi_1322550_emb_CAA96756_

(Z72576) ORF YGL054c [Saccharomyces cerevisiae]

Seq. No. 9643

Contig ID 16698_1.R1040

5'-most EST LIB3049-030-Q1-E1-G6

Seq. No. 9644

Contig ID 16703 1.R1040

5'-most EST LIB3049-030-Q1-E1-H1

Seq. No. 9645

Contig ID 16703_2.R1040 5'-most EST 2DC-01-Q1-E1-G3

Seq. No. 9646

Contig ID 16709 1.R1040

5'-most EST LIB3107-041-Q1-K1-H3

Method BLASTX
NCBI GI 94206197
BLAST score 926
E value 1.0e-100
Match length 217
% identity 12

NCBI Description (AF071527) putative pre-mRNA splicing factor [Arabidopsis

thaliana]

Seq. No. 9647

Contig ID 16710 1.R1040

5'-most EST LIB3170-031-Q1-K1-G1

Seq. No. 9648

Contig ID 16719 1.R1040

5'-most EST LIB3049-030-Q1-E1-E5

Seq. No. 9649

Contig ID 16721_1.R1040

5'-most EST jC-gmro02910040d06a1

Method BLASTX
NCBI GI g4468993
BLAST score 1496
E value 1.0e-167
Match length 347
% identity 80

NCBI Description (AL035605) putative protein [Arabidopsis thaliana]

Seq. No. 9650

Contig ID 16721 2.R1040

5'-most EST jC-gmle01810024h06a1

Method BLASTX NCBI GI g4468993



```
BLAST score
                   954
                   1.0e-103
E value
Match length
                  229
% identity
                  77
NCBI Description (AL035605) putative protein [Arabidopsis thaliana]
                  9651
Seq. No.
Contig ID
                  16727 1.R1040
5'-most EST
                  LIB3109-043-Q1-K1-F12
Method
                  BLASTX
                  g4455287
```

NCBI GI BLAST score 380 E value 2.0e-36 Match length 140 % identity 54

NCBI Description (AL035527) putative protein [Arabidopsis thaliana]

Seq. No.

16727 3.R1040 Contig ID hrw701058080.h1 5'-most EST

Seq. No. 9653

Contig ID 16732 1.R1040

5'-most EST LIB3170-030-Q1-K1-E1

Method BLASTX NCBI GI g2618689 BLAST score 545 E value 4.0e~56 Match length 108 94 % identity

NCBI Description (AC002510) unknown protein [Arabidopsis thaliana]

9654 Seq. No.

Contig ID 16746 1.R1040 5'-most EST uC-gmropic023f01b1

Method BLASTX NCBI GI q3046695 BLAST score 284 E value 1.0e-24 Match length 136 49 % identity

NCBI Description (AL022224) putative protein [Arabidopsis thaliana]

Seq. No. 9655

16746 2.R1040 Contig ID

5'-most EST LIB3056-001-Q1-B1-H4

Seq. No. 9656

Contig ID 16746 3.R1040

5'-most EST LIB3094-038-Q1-K1-F3

Method BLASTX g3046695 NCBI GI BLAST score 223 E value 7.0e-18 Match length 135 % identity 33

NCBI Description (AL022224) putative protein [Arabidopsis thaliana]



Contig ID 16753 1.R1040

5'-most EST LIB3170-031-Q1-K1-B1

Seq. No. 9658

Contig ID 16754_1.R1040

5'-most EST LIB3170-030-Q1-K1-B1

Seq. No. 9659

Contig ID 16759_1.R1040 5'-most EST gsv701051528.h1

Method BLASTX
NCBI GI g2827544
BLAST score 661
E value 4.0e-69
Match length 153
% identity 82

NCBI Description (AL021635) HSP associated protein like [Arabidopsis

thaliana]

Seq. No. 9660

Contig ID 16764 1.R1040

5'-most EST LIB3049-030-Q1-E1-A5

Method BLASTX
NCBI GI g4454026
BLAST score 283
E value 5.0e-25
Match length 76
% identity 72

NCBI Description (AL035394) phosphatase like protein [Arabidopsis thaliana]

Seq. No. 9661

Contig ID 16764 2.R1040

5'-most EST jC-gmle01810007h11d1

Method BLASTX
NCBI GI g4454026
BLAST score 268
E value 2.0e-23
Match length 60
% identity 83

NCBI Description (AL035394) phosphatase like protein [Arabidopsis thaliana]

Seq. No. 9662

Contig ID 16765_1.R1040

5'-most EST LIB $30\overline{4}$ 9-030-Q1-E1-A6

Method BLASTX
NCBI GI g2194123
BLAST score 502
E value 2.0e-50
Match length 201
% identity 54

NCBI Description (AC002062) Similar to Secale chloroplast ribosomal protein

L12 (gb SCL121A). EST gb H36579 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 9663



```
16772 1.R1040
Contig ID
                  ary700764475.hl
5'-most EST
                  BLASTX
Method
                   q4262226
NCBI GI
                   249
BLAST score
                   4.0e-21
E value
                   86
Match length
                   56
% identity
                  (AC006200) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   9664
Seq. No.
                   16772 2.R1040
Contig ID
                   jC-gmro02910007b02a1
5'-most EST
                   BLASTX
Method
                   q4262226
NCBI GI
                   244
BLAST score
                   1.0e-20
E value
                   83
Match length
                   58
% identity
                  (AC006200) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   9665
Seq. No.
                   16779 1.R1040
Contig ID
                   dpv701100971.h1
5'-most EST
                   BLASTX
Method
                   q2160180
NCBI GI
                   567
BLAST score
                   2.0e-58
E value
                   158
Match length
                   66
% identity
                   (AC000132) No definition line found [Arabidopsis thaliana]
NCBI Description
                   9666
Seq. No.
                   16783 2.R1040
Contig ID
                   LIB3139-117-P1-N1-G6
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3367576
                   515
BLAST score
                    1.0e-51
E value
                   145
Match length
                    67
 % identity
                   (AL031135) NAM / CUC2 -like protein [Arabidopsis thaliana]
NCBI Description
 Seq. No.
                    9667
                    16783 5.R1040
 Contig ID
 5'-most EST
                    fC-gmle700557414r2
                    9668
 Seq. No.
                    16783 6.R1040
 Contig ID
                    hyd700727566.h1
 5'-most EST
                    BLASTX
 Method
                    g3367576
 NCBI GI
                    781
 BLAST score
                    9.0e-83
 E value
                    475
 Match length
```

(AL031135) NAM / CUC2 -like protein [Arabidopsis thaliana]

41

% identity

NCBI Description



Contig ID 16783_9.R1040 5'-most EST ssr700554109.h1

Seq. No. 9670

Contig ID 16783_10.R1040 5'-most EST uC-gmropic054g12b1

Seq. No. 9671

Contig ID 16787 1.R1040

5'-most EST LIB3049-029-Q1-E1-H5

Seq. No. 9672

Contig ID 16789_1.R1040

5'-most EST LIB3049-029-Q1-E1-H8

Seq. No. 9673

Contig ID 16792_1.R1040

5'-most EST LIB3049-008-Q1-E1-E6

Seq. No. 9674

Contig ID 16795 1.R1040

5'-most EST uC-gmrominsoy098a04b1

Method BLASTX
NCBI GI g2252634
BLAST score 447
E value 3.0e-44
Match length 174
% identity 46

NCBI Description (U95973) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9675

Contig ID 16797_1.R1040 5'-most EST pmv700890089.h1

Seq. No. 9676

Contig ID 16798_1.R1040 5'-most EST crh700851588.h1

Seq. No. 9677

Contig ID 16803_1.R1040 5'-most EST jsh701063770.h1

Seq. No. 9678

Contig ID 16815 1.R1040 5'-most EST pcp700993566.h1

Method BLASTX
NCBI GI g1168750
BLAST score 170
E value 7.0e-12
Match length 97
% identity 40

NCBI Description CALCINEURIN B SUBUNIT (PROTEIN PHOSPHATASE 2B REGULATORY

SUBUNIT) >gi_458230 (U04380) calcineurin B [Naegleria

gruberi]



Contig ID 16815_2.R1040 5'-most EST hyd700726431.h1

Method BLASTX
NCBI GI g1168750
BLAST score 206
E value 5.0e-16
Match length 121
% identity 40

NCBI Description CALCINEURIN B SUBUNIT (PROTEIN PHOSPHATASE 2B REGULATORY

SUBUNIT) >gi_458230 (U04380) calcineurin B [Naegleria

gruberi]

Seq. No. 9680

Contig ID 16816_1.R1040 5'-most EST zzp700835437.h1

Method BLASTX
NCBI GI g2673920
BLAST score 287
E value 1.0e-25
Match length 70
% identity 80

NCBI Description (AC002561) similar to Drosophila couch potato protein

[Arabidopsis thaliana]

Seq. No. 9681

Contig ID 16817 1.R1040

5'-most EST LIB3170-020-Q1-K1-A7

Seq. No. 9682

Contig ID 16820 1.R1040

5'-most EST LIB3049-029-Q1-E1-E3

Seq. No. 9683

Contig ID 16822_1.R1040

5'-most EST jC-gmle01810009h10a1

Method BLASTX
NCBI GI g4454020
BLAST score 448
E value 3.0e-44
Match length 154
% identity 63

NCBI Description (AL035396) putative protein [Arabidopsis thaliana]

Seq. No. 9684

Contig ID 16826 1.R1040

5'-most EST LIB3049-029-Q1-E1-F1

Seq. No. 9685

Contig ID 16828_1.R1040

5'-most EST LIB3170-026-Q1-K1-B12

Method BLASTX
NCBI GI g4335764
BLAST score 443
E value 1.0e-43
Match length 207
% identity 37



NCBI Description (AC006284) putative WRKY DNA-binding protein [Arabidopsis thaliana]

Seq. No. 9686

Contig ID 16830_1.R1040 5'-most EST leu701152566.h1

Method BLASTX
NCBI GI g3482967
BLAST score 659
E value 4.0e-69
Match length 154
% identity 77

NCBI Description (AL031369) Protein phosphatase 2C-like protein [Arabidopsis thaliana] >gi_4559345_gb_AAD23006.1_AC006585_1 (AC006585)

protein phosphatase 2C [Arabidopsis thaliana]

Seq. No.

Contig ID 16830 2.R1040

5'-most EST LIB3049-029-Q1-E1-C12

9687

Method BLASTX
NCBI GI g3482967
BLAST score 456
E value 2.0e-45
Match length 114
% identity 75

NCBI Description (AL031369) Protein phosphatase 2C-like protein [Arabidopsis

thaliana] >gi 4559345 gb AAD23006.1 AC006585 1 (AC006585)

protein phosphatase 2C [Arabidopsis thaliana]

Seq. No. 9688

Contig ID 16830 4.R1040 5'-most EST kl1701213354.h1

Method BLASTX
NCBI GI g3482967
BLAST score 276
E value 2.0e-24
Match length 67
% identity 81

NCBI Description (AL031369) Protein phosphatase 2C-like protein [Arabidopsis

thaliana] >gi_4559345_gb_AAD23006.1_AC006585_1 (AC006585)

protein phosphatase 2C [Arabidopsis thaliana]

Seq. No. 9689

Contig ID 16836_1.R1040 5'-most EST leu701153534.h1

Method BLASTX
NCBI GI g4508078
BLAST score 470
E value 5.0e-47
Match length 96
% identity 89

NCBI Description (AC005882) 64134 [Arabidopsis thaliana]

Seq. No. 9690

Contig ID 16836 2.R1040

5'-most EST LIB3107-013-Q1-K1-D12

Method BLASTX



```
NCBI GI g4508078
BLAST score 210
E value 1.0e-16
Match length 45
% identity 89
NCBI Description (AC005882) 64134 [Arabidopsis thaliana]
```

Contig ID 16843 1.R1040

5'-most EST LIB3049-029-Q1-E1-A4

Seq. No. 9692

Contig ID 16844_1.R1040 5'-most EST gsv701055528.h1

Seq. No. 9693

Contig ID 16845_1.R1040

5'-most EST LIB3049-029-Q1-E1-A6

Method BLASTX
NCBI GI g4539324
BLAST score 298
E value 2.0e-37
Match length 131
% identity 68

NCBI Description (AL035679) kinesin like protein [Arabidopsis thaliana]

Seq. No. 9694

Contig ID 16846 1.R1040 5'-most EST fua701040502.h1

Seq. No. 9695

Contig ID 16847_1.R1040

5'-most EST LIB3049-029-Q1-E1-A8

Method BLASTX
NCBI GI g3953463
BLAST score 810
E value 2.0e-86
Match length 319
% identity 50

NCBI Description (AC002328) F20N2.8 [Arabidopsis thaliana]

Seq. No. 9696

Contig ID 16847_2.R1040 5'-most EST pmv700891029.h1

Method BLASTX
NCBI GI g3953463
BLAST score 138
E value 1.0e-08
Match length 34
% identity 62

NCBI Description (AC002328) F20N2.8 [Arabidopsis thaliana]

Seq. No. 9697

Contig ID 16848_1.R1040

5'-most EST LIB3109-033-Q1-K1-G4

Seq. No. 9698

16849 1.R1040 Contig ID LIB3170-025-Q1-J1-G12 5'-most EST 9699 Seq. No. 16850 1.R1040 Contig ID uC-qmflminsoy117g02b1 5'-most EST BLASTX Method NCBI GI g2832357 203 BLAST score 2.0e-15 E value 96 Match length 46 % identity (Y14071) HMG protein [Arabidopsis thaliana] >gi_3068715 NCBI Description (AF049236) unknown [Arabidopsis thaliana] Seq. No. 16850 3.R1040 Contig ID LIB3093-038-Q1-K1-H12 5'-most EST Seq. No. 16850 5.R1040 Contig ID awf700837862.h15'-most EST 9702 Seq. No. 16850 6.R1040 Contig ID wrg700791125.hl 5'-most EST

9703 Seq. No.

16851 1.R1040 Contig ID

LIB3049-029-Q1-E1-B12 5'-most EST

BLASTX Method g3033375 NCBI GI 255 BLAST score 7.0e-29 E value Match length 129 % identity 49

(AC004238) putative berberine bridge enzyme [Arabidopsis NCBI Description

thaliana]

9704 Seq. No.

16852 1.R1040 Contig ID

LIB3049-029-Q1-E1-B2 5'-most EST

BLASTX Method q4185143 NCBI GI BLAST score 244 1.0e-20 E value 93 Match length % identity

(AC005724) putative signal recognition particle receptor NCBI Description

beta subunit [Arabidopsis thaliana]

9705 Seq. No.

16854 1.R1040 Contig ID

5'-most EST LIB3049-029-Q1-E1-B5

Seq. No.

16855 1.R1040 Contig ID



zsq701118742.h1 5'-most EST

BLASTX Method g710626 NCBI GI 225 BLAST score 6.0e-19 E value 86 Match length 58 % identity

(D30719) ERD15 protein [Arabidopsis thaliana] >gi_3241941 NCBI Description

(AC004625) dehydration-induced protein ERD15 [Arabidopsis thaliana] >gi_3894181 (AC005662) ERD15 protein [Arabidopsis

thaliana]

9707 Seq. No.

16857 1.R1040 Contig ID taw700660090.hl 5'-most EST

9708 Seq. No.

16868 1.R1040 Contig ID $r1r70\overline{0}898001.h1$ 5'-most EST

9709 Seq. No.

16868 2.R1040 Contig ID

LIB3050-010-Q1-E1-B12 5'-most EST

9710 Seq. No.

16886 1.R1040 Contig ID

LIB3049-028-Q1-E1-G10 5'-most EST

BLASTX Method g3269297 NCBI GI 235 BLAST score 9.0e-20 E value 82 Match length % identity

(AL030978) putative protein [Arabidopsis thaliana] NCBI Description

9711 Seq. No.

16887 1.R1040 Contig ID

5'-most EST LIB3109-052-Q1-K1-E11

9712 Seq. No.

16887 2.R1040 Contig ID $gsv70\overline{1}045817.h1$ 5'-most EST

9713 Seq. No.

16894 1.R1040 Contig ID

5'-most EST LIB3170-028-Q1-K1-G10

BLASTX Method g1888357 NCBI GI BLAST score 186 5.0e-14 E value 41 Match length

% identity (X98130) alpha-mannosidase [Arabidopsis thaliana] NCBI Description

>gi_1890154_emb_CAA72432_ (Y11767) alpha-mannosidase

precursor [Arabidopsis thaliana]

9714 Seq. No.



```
Contig ID
                  16897 1.R1040
5'-most EST
                  sat701009333.hl
                  BLASTX
Method
                  q629623
NCBI GI
                  155
BLAST score
E value
                  3.0e-10
                  70
Match length
                  50
% identity
NCBI Description chalcone reductase - alfalfa >gi_563538_emb_CAA57784_
                  (X82368) chalcone reductase [Medicago sativa]
Seq. No.
                  9715
                  16898 1.R1040
Contig ID
                  ncj700983652.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4151319
BLAST score
                  892
E value
                  2.0e-96
Match length
                  250
% identity
                  73
NCBI Description (AF089084) putative auxin efflux carrier protein; AtPIN1
                  [Arabidopsis thaliana]
Seq. No.
                  9716
                  16900 1.R1040
Contig ID
5'-most EST
                  zzp700834260.h1
Method
                  BLASTX
                  q927428
NCBI GI
BLAST score
                  515
E value
                  3.0e-52
Match length
                  148
% identity
NCBI Description (X86733) fis1 [Linum usitatissimum]
                  9717
Seq. No.
Contig ID
                  16902 1.R1040
5'-most EST
                  LIB3051-074-Q1-K1-H2
Method
                  BLASTX
                  g2853097
NCBI GI
BLAST score
                  212
E value
                  6.0e-17
Match length
                  86
% identity
                  48
NCBI Description (AL021767) hypothetical protein [Schizosaccharomyces pombe]
Seq. No.
                  9718
                  16902 2.R1040
Contig ID
                  LIB3139-049-P1-N1-B7
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2853097
BLAST score
                  236
                  2.0e-19
E value
Match length
                  142
                  38
% identity
NCBI Description (AL021767) hypothetical protein [Schizosaccharomyces pombe]
```



Contig ID 16902 3.R1040

5'-most EST uC-gmflminsoy028c01b1

Method BLASTX
NCBI GI g1353053
BLAST score 157
E value 1.0e-10
Match length 62
% identity 47

NCBI Description HYPOTHETICAL 22.5 KD PROTEIN IN SPC1-ILV3 INTERGENIC REGION

>gi_1077868_pir__S55203 hypothetical protein YJR014w yeast (Saccharomyces cerevisiae) >gi_854588_emb_CAA60937_

(X87611) ORF YJR83.12 [Saccharomyces cerevisiae] >gi_1015646_emb_CAA89538_ (Z49514) ORF YJR014w

[Saccharomyces cerevisiae]

Seq. No. 9720

Contig ID 16909 1.R1040

5'-most EST LIB3170-026-Q1-K1-D11

Method BLASTX
NCBI GI g2632019
BLAST score 335
E value 7.0e-31
Match length 228
% identity 38

NCBI Description (AJ002571) YkfB [Bacillus subtilis]

>gi_2633652_emb_CAB13155_ (Z99110) similar to chloromuconate cycloisomerase [Bacillus subtilis]

Seq. No. 9721

Contig ID 16914_1.R1040 5'-most EST crh700853264.h1

Method BLASTN
NCBI GI g2687435
BLAST score 766
E value 0.0e+00
Match length 874
% identity 97

NCBI Description Hamamelis virginiana large subunit 26S ribosomal RNA gene,

partial sequence

Seq. No. 9722

Contig ID 16916 1.R1040

5'-most EST LIB3170-028-Q1-K1-A12

Method BLASTX
NCBI GI g4337206
BLAST score 162
E value 4.0e-11
Match length 63
% identity 54

NCBI Description (AC006403) putative replication factor-A protein

[Arabidopsis thaliana]

Seq. No. 9723

Contig ID 16918 1.R1040

5'-most EST LIB3049-028-Q1-E1-C12

Method BLASTX NCBI GI g2196466



53

113

87

6.0e-21

complete cds

BLAST score

Match length

% identity

E value

```
BLAST score
                   5.0e-60
E value
                   150
Match length
% identity
                   (Y13673) TATA binding protein-associated factor
NCBI Description
                   [Arabidopsis thaliana]
                   9724
Seq. No.
                   16919 1.R1040
Contig ID
                   LIB3049-028-Q1-E1-C3
5'-most EST
                   BLASTX
Method
                   q4115371
NCBI GI
                   829
BLAST score
                   2.0e-88
E value
                   597
Match length
                   39
% identity
                  (AC005967) unknown protein [Arabidopsis thaliana]
NCBI Description
                   9725
Seq. No.
                   16921 1.R1040
Contig ID
                   ncj700986793.hl
5'-most EST
                   BLASTX
Method
                   q2191165
NCBI GI
BLAST score
                   347
                   5.0e-32
E value
                   125
Match length
% identity
                   55
                   (AF007270) A_IG002P16.14 gene product [Arabidopsis
NCBI Description
                   thaliana]
                   9726
Seq. No.
                   16921 2.R1040
Contig ID
                   pmv700893202.h1
 5'-most EST
                   BLASTX
Method
NCBI GI
                   g2191165
BLAST score
                   329
                   3.0e - 30
 E value
                   126
Match length
 % identity
                   52
                   (AF007270) A IG002P16.14 gene product [Arabidopsis
 NCBI Description
                   thaliana]
                   9727
 Seq. No.
                    16921 4.R1040
 Contig ID
 5'-most EST
                   jC-gmst02400050c02d1
                    9728
 Seq. No.
                   16930 1.R1040
 Contig ID
                   LIB3049-013-Q1-E1-F4
 5'-most EST
                   BLASTN
 Method
                    g1747309
 NCBI GI
```

1718

NCBI Description Arabidopsis thaliana mRNA for Myb-like DNA binding protein,



```
9729
Seq. No.
                  16930 2.R1040
Contig ID
                  LIB3049-028-Q1-E1-B7
5'-most EST
Method
                  BLASTN
                  g2160155
NCBI GI
BLAST score
                  71
E value
                  1.0e-31
Match length
                  171
% identity
                   85
NCBI Description
                  Sequence of BAC F21M12 from Arabidopsis thaliana chromosome
                   1, complete sequence [Arabidopsis thaliana]
                   9730
Seq. No.
Contig ID
                  16933 1.R1040
5'-most EST
                  LIB3049-028-Q1-E1-A11
Method
                  BLASTX
NCBI GI
                   q3021513
BLAST score
                   703
E value
                   3.0e-74
Match length
                   189
% identity
                   72
                  (X96728) isocitrate dehydrogenase (NADP+) [Nicotiana
NCBI Description
                   tabacum]
                   9731
Seq. No.
                   16937 1.R1040
Contig ID
5'-most EST
                  gsv701055145.hl
Seq. No.
                   9732
                  16939 1.R1040
Contig ID
5'-most EST
                   seb70\overline{0}651907.h1
Method
                  BLASTX
NCBI GI
                   q1903021
BLAST score
                   510
E value
                   1.0e-51
Match length
                   125
                  79
% identity
                  (Y10216) hypothetical 3-isopropylmalate dehydrogenase
NCBI Description
                   [Arabidopsis thaliana]
                   9733
Seq. No.
                   16939 2.R1040
Contig ID
5'-most EST
                  LIB3049-028-Q1-E1-A8
Method
                  BLASTX
NCBI GI
                   q1903021
BLAST score
                   545
E value
                   8.0e-56
Match length
                   178
% identity
                   62
                   (Y10216) hypothetical 3-isopropylmalate dehydrogenase
NCBI Description
```

Contig ID 16939_3.R1040

5'-most EST LIB3049-015-Q1-E1-D6

[Arabidopsis thaliana]

Method BLASTX



```
a126201
NCBI GI
                     465
BLAST score
                     2.0e-46
E value
Match length
                     147
                     65
% identity
                     3-ISOPROPYLMALATE DEHYDROGENASE PRECURSOR (BETA-IPM
NCBI Description
                     DEHYDROGENASE) (IMDH) (3-IPM-DH) >gi_81676_pir__S20510 3-isopropylmalate dehydrogenase (EC 1.1.1.85) precursor -
                     rape >gi_17827_emb_CAA42596_ (X59970) 3-isopropylmalate
                     dehydrogenase [Brassica napus]
                     9735
Seq. No.
                     16942 1.R1040
Contig ID
                     kl1701203442.h1
5'-most EST
```

9736 Seq. No. 16943_1.R1040 Contig ID LIB3170-057-Q1-K1-E7 5'-most EST BLASTN Method g2924651. NCBI GI BLAST score 39 3.0e-12E value 150 Match length

 Seq. No.
 9737

 Contig ID
 16943_2.R1040

 5'-most EST
 txt700731937.h1

 Method
 BLASTN

 NCBI GI
 g2924651

 BLAST score
 35

 F value
 4.0e-10

E value 4.0e-10
Match length 59
% identity 90

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K2A18, complete sequence [Arabidopsis thaliana]

 Seq. No.
 9738

 Contig ID
 16943_3.R1040

 5'-most EST
 epx701108086.h1

 Method
 BLASTN

 NCBI GI
 g2924651

 BLAST score
 35

 F walke
 4.0e-10

BLAST score 35 E value 4.0e-10 Match length 109 % identity 90

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:

K2A18, complete sequence [Arabidopsis thaliana]

Seq. No. 9739

Contig ID 16943_4.R1040 5'-most EST uC-gmropic063e11b1

Seq. No. 9740

Contig ID 16947_1.R1040



```
jex700904806.h1
5'-most EST
                  9741
Seq. No.
                  16947 2.R1040
Contig ID
                  sat701008764.hl
5'-most EST
                   9742
Seq. No.
                   16950 1.R1040
Contig ID
5'-most EST
                  LIB3106-026-Q1-K1-G8
Seq. No.
                   9743
                   16950 2.R1040
Contig ID
5'-most EST
                   LIB3138-012-Q1-N1-D9
                   9744
Seq. No.
                   16952 1.R1040
Contig ID
                   LIB3087-004-Q1-K1-C7
5'-most EST
                   9745
Seq. No.
                   16955 1.R1040
Contig ID
                   LIB3049-027-Q1-E1-H1
5'-most EST
Seq. No.
                   9746
                   16967 1.R1040
Contig ID
5'-most EST
                   sat701002995.h1
Method
                   BLASTX
NCBI GI
                   q4049344
BLAST score
                   184
E value
                   4.0e-13
Match length
                   120
% identity
                   13
                   (AL034567) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   16967 2.R1040
Contig ID
                   LIB3049-027-Q1-E1-E7
5'-most EST
                   9748
Seq. No.
                   16970 1.R1040
Contig ID
                   LIB3049-027-Q1-E1-F1
5'-most EST
                   BLASTN
Method
                   g2656031
NCBI GI
                   34
BLAST score
                   2.0e-09
E value
Match length
                   83
                   92
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                   MXC20
                   9749
Seq. No.
                   16976 1.R1040
Contig ID
```

eep700868308.hl 5'-most EST

Method BLASTX NCBI GI g4262233 BLAST score 218 2.0e-17 E value 151 Match length



% identity 40
NCBI Description (AC006200) hypothetical protein [Arabidopsis thaliana]
Seq. No. 9750

Contig ID 16984_1.R1040

5'-most EST LIB3049-048-Q1-E1-A3

Seq. No. 9751

Contig ID 16987_1.R1040 5'-most EST LIB3049-027-Q1-E1-D4

Method BLASTX
NCBI GI g4539306
BLAST score 406
E value 3.0e-39
Match length 128
% identity 17

NCBI Description (AL049480) pumilio-like protein [Arabidopsis thaliana]

Seq. No. 9752

Contig ID 16987 2.R1040 5'-most EST sat701007693.h1

Method BLASTX
NCBI GI g4539306
BLAST score 161
E value 3.0e-11
Match length 48
% identity 62

NCBI Description (AL049480) pumilio-like protein [Arabidopsis thaliana]

Seq. No. 9753

Contig ID 16989_1.R1040

5'-most EST LIB3167-023-P4-K4-D6

Method BLASTX
NCBI GI g2924772
BLAST score 686
E value 3.0e-72
Match length 143
% identity 87

NCBI Description (AC002334) unknown protein [Arabidopsis thaliana]

Seq. No. 9754

Contig ID 16993_1.R1040 5'-most EST leu701156949.h1

Method BLASTN
NCBI GI g20755
BLAST score 475
E value 0.0e+00
Match length 691
% identity 93

NCBI Description P.sativum mRNA rab for ras-related GTP-binding protein

Seq. No. 9755

Contig ID 16993_2.R1040

5'-most EST LIB3049-027-Q1-E1-E1

Method BLASTN
NCBI GI g20755
BLAST score 173

1722



E value 2.0e-92 Match length 289 % identity 95

NCBI Description P.sativum mRNA rab for ras-related GTP-binding protein

Seq. No. 9756

Contig ID 16993_3.R1040 5'-most EST zhf700964965.h1

Method BLASTN
NCBI GI g1370185
BLAST score 55
E value 3.0e-22
Match length 107
% identity 89

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB7C

Seq. No. 9757

Contig ID 16993_4.R1040 5'-most EST rca700999277.h1

Method BLASTN
NCBI GI g20755
BLAST score 245
E value 1.0e-135
Match length 366
% identity 94

NCBI Description P.sativum mRNA rab for ras-related GTP-binding protein

Seq. No. 9758

Contig ID 16993_8.R1040 5'-most EST cf1700863676.h1

Method BLASTN
NCBI GI g1370185
BLAST score 118
E value 9.0e-60
Match length 224
% identity 92

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB7C

Seq. No. 9759

Contig ID 16997 1.R1040

5'-most EST LIB3049-027-Q1-E1-E2

Seq. No. 9760

Contig ID 17003_1.R1040 5'-most EST vwf700674904.h1

Method BLASTX
NCBI GI g2160166
BLAST score 806
E value 5.0e-86
Match length 258
% identity 60

NCBI Description (AC000132) No definition line found [Arabidopsis thaliana]

Seq. No. 9761

Contig ID 17007_1.R1040 5'-most EST uC-gmropic033f02b1

Method BLASTX



NCBI GI g4415912 BLAST score 312 E value 3.0e-29 Match length 189 % identity 38

NCBI Description (AC006282) putative protease [Arabidopsis thaliana]

Seq. No.

9762

Contig ID

17010 1.R1040

5'-most EST

LIB3049-002-Q1-E1-E1

Seq. No.

9763

Contig ID

17014 1.R1040

5'-most EST

LIB3170-040-Q1-K1-E5

Method BLASTX
NCBI GI g2245018
BLAST score 166
E value 2.0e-11
Match length 89

% identity 43

NCBI Description (Z97341) unnamed protein product [Arabidopsis thaliana]

Seq. No.

9764

Contig ID

17014 2.R1040

5'-most EST

LIB3109-015-Q1-K1-A3

Method BLASTX
NCBI GI g3776027
BLAST score 165
E value 1.0e-18
Match length 81
% identity 60

NCBI Description (AJ010475) RNA helicase [Arabidopsis thaliana]

Seq. No.

9765

Contig ID

17015_1.R1040

5'-most EST

LIB3049-027-Q1-E1-C3

Method BLASTX
NCBI GI g2583134
BLAST score 142
E value 7.0e-09
Match length 85
% identity 51

NCBI Description (AC002387) putative proline-rich protein [Arabidopsis

thaliana]

Seq. No.

9766

Contig ID 5'-most EST

17017_1.R1040 kl1701211990.h1

Seq. No.

9767

Contig ID

17019 1.R1040

5'-most EST

LIB3049-026-Q1-E1-G9

Seq. No.

9768

Contig ID

17027 1.R1040

5'-most EST

LIB3049-026-Q1-E1-H9



17027 2.R1040 Contig ID 5'-most EST q4292567

9770 Seq. No.

17030 1.R1040 Contig ID

5'-most EST LIB3049-026-Q1-E1-G8

Method BLASTX NCBI GI g2832682 BLAST score 275 E value 3.0e-24 Match length 111 53 % identity

(AL021712) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No.

Contig ID 17039 1.R1040

5'-most EST LIB3049-002-Q1-E1-A12

9772 Seq. No.

Contig ID 17041 1.R1040

5'-most EST LIB3170-027-Q2-K1-C9

Method BLASTX NCBI GI q4454465 BLAST score 161 E value 8.0e-11 Match length 95 % identity 39

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

9773 Seq. No.

Contig ID 17041 2.R1040 5'-most EST epx701106554.h1

Method BLASTX NCBI GI g4454465 BLAST score 193 E value 1.0e-14 Match length 104 % identity 40

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No. 9774

17047 1.R1040 Contig ID

5'-most EST LIB3092-012-Q1-K1-C6

9775

Method BLASTX g2529683 NCBI GI BLAST score 769 E value 1.0e-81 Match length 280 % identity 56

(AC002535) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

Contig ID 17056 1.R1040 5'-most EST vzy700755484.hl

Seq. No. 9776



Contig ID 17056_2.R1040 5'-most EST xpa700796954.h1

Seq. No. 9777

Contig ID 17058 1.R1040

5'-most EST LIB3170-057-Q1-J1-C8

Seq. No. 9778

Contig ID 17059_1.R1040 5'-most EST leu701152740.h1

Seq. No. 9779

Contig ID 17062 1.R1040

5'-most EST LIB3049-026-Q1-E1-E1

Seq. No. 9780

Contig ID 17064 1.R1040

5'-most EST jC-gmle01810037d04a2

Method BLASTX
NCBI GI g1345970
BLAST score 231
E value 4.0e-19
Match length 117
% identity 48

NCBI Description OMEGA-3 FATTY ACID DESATURASE, CHLOROPLAST PRECURSOR

>qi 870784 (U25817) omega-3 fatty acid desaturase [Sesamum

indicum]

Seq. No. 9781

Contig ID 17070 1.R1040

5'-most EST uC-gmrominsoy187a07b1

Method BLASTX
NCBI GI g4378166
BLAST score 178
E value 5.0e-13
Match length 78
% identity 47

NCBI Description (AF102543) unknown [Zymomonas mobilis]

Seq. No. 9782

Contig ID 17071_1.R1040 5'-most EST leu701157142.h1

Method BLASTX
NCBI GI g2651313
BLAST score 546
E value 1.0e-55
Match length 268
% identity 39

NCBI Description (AC002336) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9783

Contig ID 17072 1.R1040

5'-most EST LIB3170-027-Q2-K1-D9

Seq. No. 9784

Contig ID 17074_1.R1040 5'-most EST leu701157294.h1



```
Method
                  BLASTX
NCBI GI
                   g3132476
                   711
BLAST score
                   4.0e-75
E value
Match length
                  208
                   68
% identity
NCBI Description (AC003096) unknown protein [Arabidopsis thaliana]
                   9785
Seq. No.
                  17081 1.R1040
Contig ID
5'-most EST
                  LIB3170-013-Q1-K1-G11
Method
                  BLASTX
NCBI GI
                   q629561
BLAST score
                   481
E value
                   3.0e-48
Match length
                   168
% identity
                   54
                  SRG1 protein - Arabidopsis thaliana
NCBI Description
                   >gi 479047_emb_CAA55654_ (X79052) SRG1 [Arabidopsis
                  thaliana]
                   9786
Seq. No.
Contig ID
                   17085 1.R1040
5'-most EST
                  epx701105472.h1
Method
                  BLASTX
NCBI GI
                   q3834303
BLAST score
                   135
E value
                   5.0e-10
Match length
                   96
                   46
% identity
NCBI Description (AC005679) F9K20.3 [Arabidopsis thaliana]
                   9787
Seq. No.
Contig ID
                   17086_1.R1040
5'-most EST
                  LIB3049-026-Q1-E1-C5
Method
                  BLASTX
NCBI GI
                   q549638
BLAST score
                   205
E value
                   5.0e-16
Match length
                   130
% identity
                   38
                  HYPOTHETICAL 39.6 KD PROTEIN IN MTD1-NUP133 INTERGENIC
NCBI Description
                   REGION >gi_539279_pir__S38159 hypothetical protein YKR081c
                   - yeast (Saccharomyces cerevisiae) >gi 415901 emb CAA81632
                   (Z27116) ORF YKR401 [Saccharomyces cerevisiae]
                   >gi 486561 emb CAA82160 (Z28306) ORF YKR081c
                   [Saccharomyces cerevisiae]
Seq. No.
                   9788
Contig ID
                   17087 1.R1040
5'-most EST
                   uC-gmflminsoy012h05b1
Method
                   BLASTX
NCBI GI
                   g4220541
BLAST score
                   494
E value
                   6.0e-50
```

142

44

Match length % identity



```
(AL035356) Rab geranylgeranyl transferase like protein
NCBI Description
                  [Arabidopsis thaliana]
                  9789
Seq. No.
                  17092 1.R1040
Contig ID
                  uC-gmrominsoy276d01b1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3297816
                  1323
BLAST score
                  1.0e-146
E value
                  359
Match length
                  73
% identity
NCBI Description (AL031032) putative protein [Arabidopsis thaliana]
                   9790
Seq. No.
                   17092 2.R1040
Contig ID
                  LIB3107-048-Q1-K1-A11
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3297816
                   379
BLAST score
                   2.0e-36
E value
Match length
                   103
                   73
% identity
                   (AL031032) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   17092 3.R1040
Contig ID
                   uC-gmrominsoy233a03b1
5'-most EST
                   BLASTX
Method
                   g3297816
NCBI GI
                   236
BLAST score
                   1.0e-19
E value
                   89
Match length
% identity
                   (AL031032) putative protein [Arabidopsis thaliana]
NCBI Description
                   9792
Seq. No.
                   17094 1.R1040
Contig ID
                   epx701105082.h1
5'-most EST
                   BLASTX
Method
                   g927575
NCBI GI
                   323
BLAST score
                   5.0e-30
E value
                   100
Match length
                   63
% identity
                   (U12926) alpha galactosidase [Glycine max]
NCBI Description
                   9793
Seq. No.
```

Contig ID 17097_1.R1040

5'-most EST LIB3093-013-Q1-K1-E9

Seq. No. 9794

Contig ID 17098_1.R1040

5'-most EST LIB3049-026-Q1-E1-A10

Method BLASTX
NCBI GI g3702343
BLAST score 386



```
3.0e-37
E value
Match length
                   157
                   52
% identity
                   (AC005397) putative homeotic gene regulator [Arabidopsis
NCBI Description
                   thaliana]
                   9795
Seq. No.
                   17106 1.R1040
Contig ID
                   LIB3139-109-P1-N1-F5
5'-most EST
Method
                   BLASTX
                   g3212877
NCBI GI
BLAST score
                   1136
                   1.0e-124
E value
Match length
                   304
% identity
                   71
                   (AC004005) Lea-like protein [Arabidopsis thaliana]
NCBI Description
                   9796
Seq. No.
                   17106 2.R1040
Contig ID
                   kl1701207511.h1
5'-most EST
                   BLASTX
Method
                   g3212877
NCBI GI
BLAST score
                   604
E value
                   9.0e-63
Match length
                   160
% identity
                   43
                   (AC004005) Lea-like protein [Arabidopsis thaliana]
NCBI Description
                   9797
Seq. No.
                   17106 5.R1040
Contig ID
                   uC-gmropic067h01b1
5'-most EST
                   BLASTX
Method
                   g3264769
NCBI GI
                   194
BLAST score
E value
                   1.0e-14
                    49
Match length
                    78
% identity
                    (AF071894) late embryogenesis-like protein [Prunus
NCBI Description
                    armeniaca]
                    9798
Seq. No.
                    17111 1.R1040
Contig ID
                    LIB30\overline{4}9-025-Q1-E1-F8
5'-most EST
                    {\tt BLASTX}
Method
                    g4558591
NCBI GI
                    164
BLAST score
                    2.0e-11
E value
                    61
Match length
                    51
 % identity
                   (AC006555) putative beta-1,3-glucanase [Arabidopsis
NCBI Description
                    thaliana]
                    9799
 Seq. No.
 Contig ID
                    17114 1.R1040
```

5'-most EST LIB3049-001-Q1-E1-C3

Method BLASTX NCBI GI g3080439



```
BLAST score 154
E value 5.0e-10
Match length 68
% identity 44
```

NCBI Description (AL022605) putative protein [Arabidopsis thaliana]

Seq. No. 9800

Contig ID 17120 1.R1040

5'-most EST LIB3049-025-Q1-E1-G7

Method BLASTX
NCBI GI g3252868
BLAST score 284
E value 2.0e-25
Match length 78
% identity 60

NCBI Description (AF033536) putative zinc transporter [Arabidopsis thaliana]

Seq. No. 9801

Contig ID 17138_1.R1040

5'-most EST LIB3109-019-Q1-K1-F12

Seq. No. 9802

Contig ID 17140 1.R1040

5'-most EST LIB3049-025-Q1-E1-E8

Method BLASTX
NCBI GI g4204283
BLAST score 217
E value 4.0e-17
Match length 75
% identity 67

NCBI Description (AC004146) Hypothetical protein [Arabidopsis thaliana]

Seq. No. 9803

Contig ID 17141 1.R1040

5'-most EST LIB3049-025-Q1-E1-F10

Seq. No. 9804

Contig ID 17143 1.R1040

5'-most EST LIB3049-025-Q1-E1-F12

Method BLASTX
NCBI GI g2661840
BLAST score 240
E value 4.0e-20
Match length 104
% identity 63

NCBI Description (Y15430) adenosine kinase [Physcomitrella patens]

Seq. No. 9805

Contig ID 17143_2.R1040 5'-most EST gsv701050303.h1

Method BLASTX
NCBI GI g2661840
BLAST score 335
E value 1.0e-31
Match length 91
% identity 71

NCBI Description (Y15430) adenosine kinase [Physcomitrella patens]

Method

NCBI GI

BLAST score

BLASTX

283

g3687237



```
9806
Seq. No.
                  17156 1.R1040
Contig ID
                  LIB3049-025-Q1-E1-C11
5'-most EST
                  9807
Seq. No.
                  17158 1.R1040
Contig ID
5'-most EST
                  jC-gmf102220148h08a1
                  9808
Seq. No.
                  17158 3.R1040
Contig ID
                  jC-qmf102220114b06a1
5'-most EST
                   9809
Seq. No.
                  17164 1.R1040
Contig ID
                  LIB3049-025-Q1-E1-D11
5'-most EST
Method
                  BLASTX
NCBI GI
                  g3769330
BLAST score
                  1936
E value
                   0.0e+00
Match length
                   387
% identity
                  (AB015131) alpha-amylase [Phaseolus vulgaris]
NCBI Description
                   9810
Seq. No.
                   17168 1.R1040
Contig ID
                   LIB3170-025-Q1-K1-E9
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2281103
BLAST score
                   239
E value
                   8.0e-20
Match length
                   115
                   42
% identity
                   (AC002333) Glucan endo-1,3-beta glucosidase isolog
NCBI Description
                   [Arabidopsis thaliana]
                   9811
Seq. No.
                   17170 1.R1040
Contig ID
5'-most EST
                   LIB3049-013-Q1-E1-D5
                   9812
Seq. No.
                   17173 1.R1040
Contig ID
                   LIB3106-026-Q1-K1-E3
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2649239
BLAST score
                   168
                   1.0e-11
E value
                   47
Match length
% identity
                  (AE001011) rubredoxin (rd-2) [Archaeoglobus fulgidus]
NCBI Description
                   9813
Seq. No.
                   17175 1.R1040
Contig ID
                   LIB3170-028-Q1-K1-F9
5'-most EST
```

1731



E value 8.0e-25 Match length 120 % identity 58

NCBI Description (AC005169) putative Cys3His zinc-finger protein

[Arabidopsis thaliana]

Seq. No. 9814

Contig ID 17175_2.R1040 5'-most EST crh700849930.h1

Method BLASTX
NCBI GI g3687237
BLAST score 352
E value 2.0e-33
Match length 136
% identity 59

NCBI Description (AC005169) putative Cys3His zinc-finger protein

[Arabidopsis thaliana]

Seq. No. 9815

Contig ID 17177 1.R1040

5'-most EST LIB3049-025-Q1-E1-A8

Method BLASTX
NCBI GI g1396054
BLAST score 218
E value 1.0e-17
Match length 50

Match length 50 % identity 76

NCBI Description (D86180) phosphoribosylanthranilate transferase [Pisum

sativum]

Seq. No.

Contig ID 17180_1.R1040 5'-most EST jex700909029.h1

Seq. No.

9817

9816

Contig ID 17180 2.R1040

5'-most EST LIB30 $\overline{4}$ 9-024-Q1-E1-H6

Seq. No.

9818

Contig ID 17185_1.R1040 5'-most EST pcp700993957.h1

Seq. No.

9819

Contig ID 17186_1.R1040

5'-most EST LIB3051-096-Q1-K1-A7

Method BLASTX
NCBI GI g1362162
BLAST score 1869
E value 0.0e+00
Match length 501
% identity 68

NCBI Description beta-glucosidase BGQ60 precursor - barley >qi 804656

(L41869) beta-glucosidase [Hordeum vulgare]

Seq. No.

9820

Contig ID

17186 2.R1040

5'-most EST LIB3051-111-Q1-K1-B10



Method BLASTX
NCBI GI g1362162
BLAST score 382
E value 6.0e-37
Match length 103
% identity 64

NCBI Description beta-glucosidase BGQ60 precursor - barley >gi_804656

(L41869) beta-glucosidase [Hordeum vulgare]

Seq. No. 9821

Contig ID 17187_1.R1040 5'-most EST gsv701055639.h1

Method BLASTX
NCBI GI g3776559
BLAST score 692
E value 5.0e-73
Match length 157
% identity 83

NCBI Description (AC005388) Strong similarity to gene F14J9.26 gi_3482933

cdc2 protein kinase homolog from A. thaliana BAC

gb_AC003970. ESTs gb_Z35332 and gb_F19907 come from this

gene. [Arabidopsis thaliana]

Seq. No. 9822

Contig ID 17192_1.R1040

5'-most EST LIB3138-079-P1-N1-H12

Method BLASTX
NCBI GI g2191187
BLAST score 175
E value 1.0e-12
Match length 95
% identity 43

NCBI Description (AF007271) contains similarity to a DNAJ-like domain

[Arabidopsis thaliana]

Seq. No. 9823

Contig ID 17194_1.R1040

5'-most EST LIB3170-026-Q1-J1-A8

Seq. No. 9824

Contig ID 17199_1.R1040

5'-most EST jC-gmst02400031h09d1

Method BLASTX
NCBI GI g2760844
BLAST score 247
E value 6.0e-21
Match length 68
% identity 68

NCBI Description (AC003105) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9825

Contig ID 17210_1.R1040

5'-most EST LIB3109-053-Q1-K1-H10

Method BLASTX
NCBI GI g4406764
BLAST score 494
E value 6.0e-51



Match length 157 % identity 36

NCBI Description (AC006836) putative uridylyl transferase [Arabidopsis

thaliana]

Seq. No. 9826

Contig ID 17210_2.R1040

5'-most EST LIB3049-024-Q1-E1-F5

Method BLASTX
NCBI GI g4406764
BLAST score 210
E value 9.0e-17
Match length 55
% identity 71

NCBI Description (AC006836) putative uridylyl transferase [Arabidopsis

thaliana]

Seq. No. 9827

Contig ID 17212 1.R1040

5'-most EST LIB3049-024-Q1-E1-E10

Method BLASTX
NCBI GI g2995953
BLAST score 294
E value 2.0e-26
Match length 92
% identity 61

NCBI Description (AF053565) glutaredoxin I [Mesembryanthemum crystallinum]

Seq. No. 9828

Contig ID 17224_1.R1040

5'-most EST jC-gmst02400057e05a1

Method BLASTX
NCBI GI g4101564
BLAST score 287
E value 2.0e-25
Match length 112
% identity 56

NCBI Description (AF004556) IFA-binding protein [Arabidopsis thaliana]

Seq. No. 9829

Contig ID 17227_1.R1040

5'-most EST LIB3049-024-Q1-E1-D8

Method BLASTX
NCBI GI g2465925
BLAST score 163
E value 2.0e-11
Match length 91
% identity 22

NCBI Description (AF024649) receptor-like serine/threonine kinase

[Arabidopsis thaliana]

Seq. No. 9830

Contig ID 17230_1.R1040

5'-most EST LIB3049-024-Q1-E1-C11

Seq. No. 9831

Contig ID 17231_1.R1040



```
epx701108947.h1
5'-most EST
                   9832
Seq. No.
                   17233 1.R1040
Contig ID
                   uxk70\overline{0}669925.h1
5'-most EST
                   BLASTX
Method
                   g1175102
NCBI GI
BLAST score
                   217
E value
                   3.0e-17
Match length
                   187
                   37
% identity
                   HYPOTHETICAL PROTEIN HI0077 >gi 1074247 pir G64000
NCBI Description
                   hypothetical protein HI0077 - Haemophilus influenzae
                   (strain Rd KW20) >gi_1573032 (U32693) H. influenzae
                   predicted coding region HI0077 [Haemophilus influenzae Rd]
                   9833
Seq. No.
                   17243 1.R1040
Contig ID
                   LIB3170-026-Q1-K1-F7
5'-most EST
                   9834
Seq. No.
                   17244 1.R1040
Contig ID
                   LIB3049-024-Q1-E1-B6
5'-most EST
                   9835
Seq. No.
                   17248 1.R1040
Contig ID
                   LIB3170-026-Q1-K1-G7
5'-most EST
                   9836
Seq. No.
                   17259 1.R1040
Contig ID
                   LIB30\overline{5}1-107-Q1-K1-D3
5'-most EST
                   BLASTX
Method
                   g1899188
NCBI GI
                   301
BLAST score
                   2.0e-27
E value
                   127
Match length
                   37
% identity
                   (U90212) DNA binding protein ACBF [Nicotiana tabacum]
NCBI Description
                   9837
Seq. No.
                   17259 2.R1040
Contig ID
                   LIB3087-009-Q1-K1-E2
5'-most EST
                   BLASTX
Method
                    q4455223
NCBI GI
BLAST score
                    396
                    2.0e-38
E value
                    116
Match length
                    23
 % identity
                    (AL035440) putative DNA binding protein [Arabidopsis
NCBI Description
                    thaliana]
```

Contig ID 17259_3.R1040

5'-most EST jC-gmro02910003f05a1

Method BLASTX
NCBI GI g2708532
BLAST score 358



7.0e-51 E value Match length 159 33 % identity

(AF029351) putative RNA binding protein [Nicotiana tabacum] NCBI Description

Seq. No.

17259 4.R1040 Contig ID 5'-most EST zhf700963840.hl

Method BLASTX g1899188 NCBI GI BLAST score 539 4.0e-55 E value 157 Match length % identity

(U90212) DNA binding protein ACBF [Nicotiana tabacum] NCBI Description

9840 Seq. No.

17265 1.R1040 Contig ID uC-gmropic103d05b1 5'-most EST

Method BLASTX NCBI GI g3023946 BLAST score 195 E value 6.0e-15 Match length 100 46 % identity

HIRA PROTEIN (TUP1 LIKE ENHANCER OF SPLIT PROTEIN 1) NCBI Description

>gi 2352031 (U94324) Tuple1/HirA [Fugu rubripes] >gi 2352036 (U94325) Tuple1/HirA [Fugu rubripes]

9841 Seq. No.

17266 1.R1040 Contig ID hrw701058819.h1 5'-most EST

9842 Seq. No.

17268 1.R1040 Contig ID bth700849208.h1 5'-most EST

BLASTX Method g2462762 NCBI GI 293 BLAST score 2.0e-26 E value 113 Match length 54 % identity

(AC002292) Highly similar to auxin-induced protein NCBI Description

(aldo/keto reductase family) [Arabidopsis thaliana]

Seq. No. 9843

17271 1.R1040 Contig ID bth700848656.hl 5'-most EST

9844 Seq. No.

17271 2.R1040 Contig ID gsv701046787.hl 5'-most EST

Seq. No.

Contig ID 17274 1.R1040 5'-most EST uC-qmropic024b07b1

BLASTX Method



```
q1632831
NCBI GI
                   627
BLAST score
                   3.0e-65
E value
Match length
                   142
                   85
% identity
                   (Z49698) orf [Ricinus communis]
NCBI Description
Seq. No.
Contig ID
                   17274 2.R1040
                   LIB3049-012-Q1-E1-H6
5'-most EST
Method
                   BLASTX
                   q4512684
NCBI GI
                   422
BLAST score
                   2.0e-41
E value
Match length
                   94
                   89
% identity
                   (AC006931) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4559324_gb_AAD22986.1_AC007087_5 (AC007087) unknown
                   protein [Arabidopsis thaliana]
                   9847
Seq. No.
                   17276 1.R1040
Contig ID
5'-most EST
                   leu701149782.h1
                   BLASTX
Method
                   q4097547
NCBI GI
BLAST score
                   277
                   2.0e-24
E value
                   68
Match length
                   43
% identity
                   (U64906) ATFP3 [Arabidopsis thaliana]
NCBI Description
                   9848
Seq. No.
                   17276 2.R1040
Contig ID
                   LIB3073-008-Q1-K1-G7
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4097547
BLAST score
                   296
                   1.0e-26
E value
Match length
                   69
                   40
% identity
                   (U64906) ATFP3 [Arabidopsis thaliana]
NCBI Description
                   9849
Seq. No.
                   17278 1.R1040
Contig ID
                   LIB3049-023-Q1-E1-E12
 5'-most EST
Method
                   BLASTX
NCBI GI
                   g2737973
BLAST score
                   460
                   7.0e-46
E value
                   146
Match length
                   66
 % identity
                   (U83625) protein kinase ZmMEK1 [Zea mays]
NCBI Description
                   9850
 Seq. No.
```

17290 1.R1040

jC-gmle01810009d08a1

Contig ID 5'-most EST



```
9851
Seq. No.
                  17290 2.R1040
Contig ID
                  LIB3049-023-Q1-E1-A10
5'-most EST
Method
                  BLASTX
                  g1345666
NCBI GI
                  216
BLAST score
                  4.0e-17
E value
                  86
Match length
                  48
% identity
                  F-ACTIN CAPPING PROTEIN BETA SUBUNIT >gi_1016279 (U35240)
NCBI Description
                  capping protein beta [Drosophila melanogaster]
                   9852
Seq. No.
                   17291 1.R1040
Contig ID
                   iC-qmst02400063b11a1
5'-most EST
                   BLASTX
Method
                   q3643611
NCBI GI
BLAST score
                   1243
                   1.0e-137
E value
Match length
                   298
                   76
% identity
                   (AC005395) putative casein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   17291 2.R1040
Contig ID
5'-most EST
                   q5126468
                   BLASTN
Method
                   q4220640
NCBI GI
BLAST score
                   33
                   8.0e-09
E value
Match length
                   61
                   89
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:
NCBI Description
                   MPE11, complete sequence [Arabidopsis thaliana]
                   9854
Seq. No.
                   17291 3.R1040
Contig ID
                   uC-gmropic014g08b1
5'-most EST
                   9855
Seq. No.
                   17291 7.R1040
Contig ID
                   uxk700672101.h1
 5'-most EST
                   BLASTX
Method
                   g3643607
NCBI GI
BLAST score
                   347
                   5.0e-33
E value
                   99
Match length
 % identity
                   (AC005395) unknown protein [Arabidopsis thaliana]
 NCBI Description
                   9856
 Seq. No.
                   17291 8.R1040
 Contig ID
```

Contig ID 17291_8.R1040
5'-most EST bth700844667.h1
Method BLASTX
NCBI GI g3643611

BLAST score 121 E value 2.0e-15



```
Match length
                  57
% identity
                  (AC005395) putative casein kinase [Arabidopsis thaliana]
NCBI Description
                  9857
Seq. No.
                  17294 1.R1040
Contig ID
                  LIB3093-006-Q1-K1-H1
5'-most EST
                  BLASTN
Method
                  g3821780
NCBI GI
BLAST score
                  36
                  2.0e-10
E value
Match length
                  36
                  100
% identity
```

NCBI Description Xenopus laevis cDNA clone 27A6-1

9858

17294 2.R1040 Contig ID 5'-most EST uaw700662762.h1 Method BLASTX NCBI GI q1651934 BLAST score 242 E value 4.0e-20 Match length 183 % identity 35

Seq. No.

NCBI Description (D90901) hypothetical protein [Synechocystis sp.]

Seq. No. 9859 Contig ID 17296 1.R1040 5'-most EST zzp700833170.h1 Method BLASTX q2505940 NCBI GI BLAST score 1181

E value 1.0e-130 Match length 309 % identity

NCBI Description (Y13071) 26S proteasome, non-ATPase subunit [Mus musculus]

9860 Seq. No.

Contig ID 17297 1.R1040 5'-most EST hyd700726184.h1

Method BLASTN g3849811 NCBI GI BLAST score 33 E value 6.0e-09 Match length 61 89 % identity

NCBI Description Arabidopsis thaliana chromosome I BAC T2P11 genomic sequence, complete sequence [Arabidopsis thaliana]

9861 Seq. No.

17306 1.R1040 Contig ID

5'-most EST LIB3051-042-Q1-K1-A9

Method BLASTX NCBI GI g2865175 BLAST score 230 E value 1.0e-18 Match length 116



% identity 40

(AB010945) AtRer1A [Arabidopsis thaliana] NCBI Description

9862 Seq. No.

Contig ID 17306 4.R1040 5'-most EST 6HA-02-Q1-B1-E11

Seq. No. 9863

Contig ID 17307 1.R1040

LIB3170-025-Q1-J1-D7 5'-most EST

Method BLASTN NCBI GI q4097568 BLAST score 84 E value 2.0e-39 Match length 132 % identity 91

NCBI Description Glycine max farnesylated protein GMFP4 mRNA, partial cds

Seq. No.

Contig ID 17308 1.R1040

5'-most EST LIB3049-023-Q1-E1-C2

9864

Method BLASTX q2494238 NCBI GI BLAST score 353 E value 3.0e-33 Match length 184 % identity 41

NCBI Description PROBABLE PEROXISOMAL ENOYL-COA HYDRATASE

> >gi 2135896 pir I38882 peroxisomal enoyl-CoA hydratase-like protein - human >gi 564065 (U16660)

peroxisomal enoyl-CoA hydratase-like protein [Homo sapiens] >gi_2623168 (AF030249) putative dienoyl-CoA isomerase [Homo sapiens] >gi 4503447 ref NP 001389.1 pECH1 enoyl Coenzyme

A hydratase 1, peroxisomal

Seq. No. 9865

17319 1.R1040 Contig ID

5'-most EST LIB3170-028-Q1-K1-H6

Method BLASTX NCBI GI g2497542 BLAST score 232 5.0e-19 E value Match length 47 91 % identity

NCBI Description PYRUVATE KINASE, CHLOROPLAST ISOZYME G PRECURSOR

>gi_629696_pir__S44287 pyruvate kinase, plastid - common tobacco >gi 482938 emb CAA82223 (Z28374) Pyruvate kinase;

plastid isozyme [Nīcotīana tabacum]

9866 Seq. No.

17327_1.R1040 Contig ID

5'-most EST LIB3170-026-Q1-K1-A6

9867 Seq. No.

Contig ID 17341 1.R1040

5'-most EST LIB3049-022-Q1-E1-F7



17343 1.R1040 Contig ID

5'-most EST jC-gmf102220065g12d1

Seq. No.

9869

17344 1.R1040 Contig ID 5'-most EST gsv701054265.hl

Seq. No.

9870

17344 3.R1040 Contig ID 5'-most EST hyd700728428.h1

Seq. No.

9871

60

Contig ID

17346 1.R1040

5'-most EST

uC-gmrominsoy107d06b1

Method BLASTX NCBI GI g120941 BLAST score 336 E value 3.0e-31 97

Match length % identity

NCBI Description

GAR1 PROTEIN >gi 83030 pir S19634 nucleolar protein GAR1 -

yeast (Saccharomyces cerevisiae) >gi 3728 emb CAA45162 (X63617) GAR1 [Saccharomyces cerevisiae] >gi 487935 (U00060) Garlp: Small nucleolar RNA protein required for

pre-rRNA splicing [Saccharomyces cerevisiae]

Seq. No. 9872

17346 2.R1040 Contig ID 5'-most EST asn701133850.h1

Method BLASTX NCBI GI q120941 BLAST score 160 7.0e-11 E value Match length 44 61

% identity

GAR1 PROTEIN >gi_83030_pir S19634 nucleolar protein GAR1 -NCBI Description

yeast (Saccharomyces cerevisiae) >gi 3728 emb CAA45162 (X63617) GAR1 [Saccharomyces cerevisiae] >gi 487935 (U00060) Garlp: Small nucleolar RNA protein required for

pre-rRNA splicing [Saccharomyces cerevisiae]

Seq. No.

9873

Contig ID

17349 1.R1040

5'-most EST LIB3050-022-Q1-K1-B1

Seq. No.

9874

Contig ID 5'-most EST

17349 2.R1040 uxk700667230.hl

Seq. No.

9875

Contig ID 5'-most EST

17349 3.R1040 $crh70\overline{0}853991.h1$

Seq. No.

9876

Contig ID

17349 4.R1040

5'-most EST

uxk700668251,h1



Contig ID 17353 1.R1040 5'-most EST hvd700729857.hl

Method BLASTX NCBI GI g112785 BLAST score 395 E value 5.0e-38 Match length 178 42 % identity

NCBI Description DNA-3-METHYLADENINE GLYCOSIDASE I (3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I) >gi_67508_pir__DGECM1

3-methyladenine DNA glycosylase (EC 3.2.2.-) I-

Escherichia coli >gi_43030_emb_CAA27472_ (X03845) TAGI (aa 1-187) [Escherichia coli] >gi_147920 (J02606)

3-methyladenine-DNA glycosylase I (tag) [Escherichia coli] >gi 466687 (U00039) 3-methyladenine DNA glycosylase I, constitutive [Escherichia coli] >gi_1789971 (AE000432)

3-methyl-adenine DNA glycosylase I, constitutive

[Escherichia coli]

Seq. No. 9878

17355 1.R1040 Contig ID fua701040371.h1 5'-most EST

9879 Seq. No.

Contig ID 17355 2.R1040 5'-most EST jsh701067431.h1

9880 Seq. No.

17356 1.R1040 Contig ID

5'-most EST LIB3065-019-Q1-N1-A5

Method BLASTX NCBI GI g3901012 BLAST score 551 E value 1.0e-113 Match length 276 % identity 72

NCBI Description (AJ130885) xyloglucan endotransglycosylase 1 [Fagus

sylvatica]

Seq. No. 9881

Contig ID 17359 1.R1040

5'-most EST LIB3049-022-Q1-E1-C8

9882 Seq. No.

17364 1.R1040 Contig ID

5'-most EST LIB3049-022-Q1-E1-B12

Method BLASTX NCBI GI q1703190 BLAST score 159 E value 8.0e-11 Match length 105 % identity 33

NCBI Description AF-10 PROTEIN >gi 538277 (U13948) zinc finger/leucine

zipper protein [Homo sapiens]



```
Seq. No.
                   9883
Contig ID
```

17366 1.R1040 LIB3049-022-Q1-E1-B3 5'-most EST

72

Method BLASTX NCBI GI g2252631 BLAST score 478 7.0e-48 E value 145 Match length

% identity NCBI Description (U95973) hypothetical protein [Arabidopsis thaliana]

Seq. No.

17366 2.R1040 Contig ID zhf700952405.h1 5'-most EST

9885 Seq. No.

17368_1.R1040 Contig ID 5'-most EST ncj700977474.hl

9886 Seq. No.

17373 1.R1040 Contig ID 5'-most EST asn701132787.h1

Method BLASTX NCBI GI g2827139 BLAST score 3734 0.0e+00 E value Match length 779 88 % identity

NCBI Description (AF027172) cellulose synthase catalytic subunit

> [Arabidopsis thaliana] >qi 4049343 emb CAA22568 (AL034567) cellulose synthase catalytic subunit (RSW1) [Arabidopsis

thaliana]

9887 Seq. No.

Contig ID 17373 2.R1040

5'-most EST jC-gmst02400003a09d1

Method BLASTX q2827143 NCBI GI BLAST score 157 E value 2.0e-10 Match length 42 74 % identity

NCBI Description (AF027174) cellulose synthase catalytic subunit

[Arabidopsis thaliana]

9888 Seq. No.

17374 1.R1040 Contig ID 5'-most EST wvk700686102.h1

Method BLASTN q791097 NCBI GI BLAST score 41 E value 7.0e-14 Match length 77 % identity 88

NCBI Description P.vulgaris plsB mRNA

9889 Seq. No.



Contig ID 17375_1.R1040

5'-most EST uC-gmrominsoy157h02b1

Method BLASTX
NCBI GI 9481762
BLAST score 162
E value 7.0e-11
Match length 181
% identity 28

NCBI Description beta-adaptin 1 - fruit fly (Drosophila melanogaster)

>gi_434902_emb_CAA53509_ (X75910) beta-adaptin Drosophila 1

[Drosophila melanogaster]

Seq. No. 9890

Contig ID 17377_1.R1040 5'-most EST leu701144426.h1

Method BLASTX
NCBI GI g3746059
BLAST score 556
E value 4.0e-57
Match length 151
% identity 65

NCBI Description (AC005311) putative cysteinyl-tRNA synthetase [Arabidopsis

thaliana] >gi_4432812_gb_AAD20662_ (AC006593) putative

cysteinyl-tRNA synthetase [Arabidopsis thaliana]

Seq. No. 9891

Contig ID 17377 2.R1040 5'-most EST leu701152494.h1

Method BLASTX
NCBI GI g3746059
BLAST score 237
E value 1.0e-19
Match length 105
% identity 49

NCBI Description (AC005311) putative cysteinyl-tRNA synthetase [Arabidopsis

thaliana] >gi_4432812_gb_AAD20662_ (AC006593) putative

cysteinyl-tRNA synthetase [Arabidopsis thaliana]

Seq. No. 9892

Contig ID 17381 1.R1040

5'-most EST LIB3049-021-Q1-E1-H5

Seq. No. 9893

Contig ID 17384 1.R1040

5'-most EST LIB3049-021-Q1-E1-H8

Seq. No. 9894

Contig ID 17385 1.R1040 5'-most EST ncj700978967.h1

Seq. No. 9895

Contig ID 17390 1.R1040

5'-most EST LIB3170-028-Q1-K1-C5

Seq. No. 9896

Contig ID 17391_1.R1040

5'-most EST LIB3049-022-Q1-E1-A3



17394 1.R1040 Contig ID zhf700953342.h1 5'-most EST

9898 Seq. No.

17399 1.R1040 Contig ID

5'-most EST jC-gmle01810084a09a1

Method BLASTX NCBI GI q121083 BLAST score 3312 E value 0.0e + 00Match length 693 % identity

NCBI Description GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR (GLYCINE

DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN) >gi_282926_pir_A42109 glycine dehydrogenase

(decarboxylating) (EC 1.4.4.2) component P precursor garden pea >gi 20741 emb CAA42443 (X59773) component of

aminomethyltransferase [Pisum satīvum]

Seq. No. 9899

17399 2.R1040 Contig ID

5'-most EST LIB3138-016-Q1-N1-D8

Method BLASTN NCBI GI q20738 BLAST score 69 E value 3.0e-30 Match length 97 % identity 93

NCBI Description P.sativum gdcP gene for P-protein of glycine decarboxylase

enzyme complex

9900 Seq. No.

Contig ID 17399 3.R1040 5'-most EST wrg700788492.h1

Method BLASTN NCBI GI q20738 BLAST score 197 E value 1.0e-106 Match length 364 % identity 91

NCBI Description P.sativum gdcP gene for P-protein of glycine decarboxylase

enzyme complex

Seq. No. 9901

Contig ID 17399 4.R1040

5'-most EST g5753163 Method BLASTN NCBI GI g20740 BLAST score 154 E value 9.0e-81 Match length 634 % identity 89

NCBI Description Pisum sativum mRNA for P protein, a part of glycine

cleavage complex



Contig ID 17407_1.R1040 5'-most EST pcp700994349.h1

Method BLASTX
NCBI GI g3880625
BLAST score 346
E value 4.0e-32
Match length 146
% identity 49

NCBI Description (Z93785) predicted using Genefinder; similar to RNA

recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL:T01682 comes from this gene; cDNA EST EMBL:M75823 comes from this gene; cDNA EST EMBL:D27559 comes from this

ge

Seq. No. 9903

Contig ID 17407_2.R1040

5'-most EST uC-gmflminsoy042h09b1

Method BLASTX
NCBI GI g3880625
BLAST score 317
E value 6.0e-29
Match length 105
% identity 57

NCBI Description (Z93785) predicted using Genefinder; similar to RNA

recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL:T01682 comes from this gene; cDNA EST EMBL:M75823 comes from this gene; cDNA EST EMBL:D27559 comes from this

ge

Seq. No. 9904

Contig ID 17407_3.R1040

5'-most EST jC-gmro02910047b12a1

Method BLASTX
NCBI GI g3880625
BLAST score 262
E value 7.0e-23
Match length 85
% identity 58

NCBI Description (Z93785) predicted using Genefinder; similar to RNA

recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL:T01682 comes from this gene; cDNA EST EMBL:M75823 comes from this gene; cDNA EST EMBL:D27559 comes from this

ge

Seq. No. 9905

Contig ID 17407_4.R1040

5'-most EST LIB3049-021-Q1-E1-G5

Seq. No. 9906

Contig ID 17407_5.R1040 5'-most EST zhf700961647.h1

Method BLASTX
NCBI GI g3880625
BLAST score 167
E value 1.0e-11
Match length 54



% identity 57

NCBI Description (Z93785) predicted using Genefinder; similar to RNA

recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL:T01682 comes from this gene; cDNA EST EMBL:M75823 comes from this gene; cDNA EST EMBL:D27559 comes from this

ge

Seq. No. 9907

Contig ID 17411_1.R1040 5'-most EST eep700868642.h1

Method BLASTX
NCBI GI g3250695
BLAST score 375
E value 7.0e-36
Match length 115
% identity 59

NCBI Description (AL024486) putative protein [Arabidopsis thaliana]

Seq. No. 9908

Contig ID 17411 2.R1040 5'-most EST dpv701100557.h1

Method BLASTX
NCBI GI g3250695
BLAST score 220
E value 1.0e-17
Match length 82
% identity 52

NCBI Description (AL024486) putative protein [Arabidopsis thaliana]

Seq. No. 9909

Contig ID 17417_1.R1040 5'-most EST uC-gmropic110c07b1

Method BLASTX
NCBI GI g3953471
BLAST score 1156
E value 1.0e-127
Match length 423

NCBI Description (AC002328) F2202.16 [Arabidopsis thaliana]

Seq. No. 9910

% identity

Contig ID 17418_1.R1040

5'-most EST LIB3051-008-Q1-E1-E8

54

Method BLASTX
NCBI GI g2961388
BLAST score 263
E value 7.0e-23
Match length 75
% identity 68

NCBI Description (AL022141) putative protein [Arabidopsis thaliana]

Seq. No. 9911

Contig ID 17423_1.R1040

5'-most EST LIB3170-028-Q1-K1-H4

Seq. No. 9912

Contig ID 17428 1.R1040



```
LIB3049-021-Q1-E1-E8
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2642450
                  452
BLAST score
                  7.0e-45
E value
Match length
                  121
                  74
% identity
                  (AC002391) putative metal ion transporter (Nramp)
NCBI Description
                   [Arabidopsis thaliana] >gi_3169188 (AC004401) putative
                  metal ion transporter (Nramp) [Arabidopsis thaliana]
                  9913
Seq. No.
                  17436 1.R1040
Contig ID
                  LIB3170-026-Q1-K1-H3
5'-most EST
                   9914
Seq. No.
                  17437 1.R1040
Contig ID
                  awf700840951.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                   g2088647
BLAST score
                   655
                   1.0e-68
E value
Match length
                   171
                   74
% identity
                   (AF002109) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   >gi_3158394 (AF036340) LRR-containing F-box protein
                   [Arabidopsis thaliana]
                   9915
Seq. No.
                   17442 1.R1040
Contig ID
                   pcp700991595.hl
5'-most EST
                   9916
Seq. No.
                   17448 1.R1040
Contig ID
                   LIB3049-008-Q1-E1-A11
5'-most EST
Method
                   BLASTX
                   q2500717
NCBI GI
                   266
BLAST score
                   3.0e-23
E value
                   124
Match length
                   49
 % identity
                   STS14 PROTEIN PRECURSOR >gi_2129995_pir__S65052
NCBI Description
                   pistil-specific protein sts14 precursor - potato
                   >gi_1236785_emb_CAA57976_ (X82652) sts14 [Solanum
                   tuberosum] >gi 1589691 prf 2211417A sts14 gene [Solanum
                   tuberosum]
                   9917
 Seq. No.
                   17453 1.R1040
```

Contig ID

LIB3170-026-Q1-K1-F3 5'-most EST

9918 Seq. No.

17456 1.R1040 Contig ID

5'-most EST LIB3049-021-Q1-E1-A4

Seq. No.

17457 1.R1040 Contig ID

% identity

62



```
bth700845843.h1
5'-most EST
Method
                  BLASTX
                  g2995198
NCBI GI
                  225
BLAST score
E value
                  3.0e-18
Match length
                  124
                  37
% identity
NCBI Description (Y11588) apoptosis specific protein [Homo sapiens]
                  9920
Seq. No.
Contig ID
                  17459 1.R1040
5'-most EST
                  LIB3106-074-Q1-K1-D5
Method
                  BLASTX
NCBI GI
                  q4098517
BLAST score
                  771
E value
                  5.0e-82
Match length
                  180
                  81
% identity
NCBI Description (U79114) auxin-binding protein ABP19 [Prunus persica]
Seq. No.
                  9921
                  17460 1.R1040
Contig ID
                  smc700750129.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3170178
BLAST score
                  175
E value
                  2.0e-12
Match length
                  95
% identity
                  40
NCBI Description (AF039689) antigen NY-CO-7 [Homo sapiens]
Seq. No.
                  9922
                  17460 2.R1040
Contig ID
5'-most EST
                  LIB3049-021-Q1-E1-A9
                   9923
Seq. No.
Contig ID
                  17462 1.R1040
5'-most EST
                  LIB3050-013-Q1-E1-A1
Method
                  BLASTX
NCBI GI
                  g2288986
BLAST score
                  710
E value
                  8.0e-75
Match length
                  191
                  70
% identity
NCBI Description
                  (AC002335) glyoxalase II isolog [Arabidopsis thaliana]
Seq. No.
                   9924
                  17462 2.R1040
Contig ID
5'-most EST
                  q5677233
Method
                  BLASTX
NCBI GI
                  g2570338
BLAST score
                  325
E value
                  6.0e-30
Match length
                  101
```

NCBI Description (U90927) glyoxalase II isozyme [Arabidopsis thaliana]



```
9925
Seq. No.
                  17462 3.R1040
Contig ID
                  LIB3055-011-Q1-N1-F3
5'-most EST
Method
                  BLASTX
                   g2570338
NCBI GI
BLAST score
                   767
                   1.0e-81
E value
                   221
Match length
% identity
                   65
                   (U90927) glyoxalase II isozyme [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   17462 4.R1040
Contig ID
                   uC-gmrominsoy170e07b1
5'-most EST
Method
                   BLASTX
                   g2288986
NCBI GI
BLAST score
                   211
                   1.0e-16
E value
                   60
Match length
                   65
% identity
                   (AC002335) glyoxalase II isolog [Arabidopsis thaliana]
NCBI Description
                   9927
Seq. No.
                   17462 7.R1040
Contig ID
                   LIB3049-021-Q1-E1-B10
5'-most EST
                   BLASTX
Method
                   q2570338
NCBI GI
                   185
BLAST score
                   5.0e-14
E value
Match length
                   56
                   57
% identity
                   (U90927) glyoxalase II isozyme [Arabidopsis thaliana]
NCBI Description
                   9928
Seq. No.
                   17465 1.R1040
Contig ID
                   gsv701047950.hl
 5'-most EST
                   9929
 Seq. No.
                   17472 1.R1040
 Contig ID
                   ssr700560174.hl
 5'-most EST
                    9930
 Seq. No.
                    17473 1.R1040
 Contig ID
                   LIB3049-020-Q1-E1-G8
 5'-most EST
                   BLASTX
 Method
                   g3643603
 NCBI GI
                    398
 BLAST score
                    9.0e-39
 E value
                    134
 Match length
 % identity
                    53
                   (AC005395) unknown protein [Arabidopsis thaliana]
 NCBI Description
                    9931
 Seq. No.
```

Contig ID 17476_1.R1040

5'-most EST LIB3051-061-Q1-K1-C10

Method BLASTX NCBI GI g2952433



BLAST score 453 E value 5.0e-45 Match length 119 % identity 76

NCBI Description (AF051135) putative ubiquitin activating enzyme E1

[Arabidopsis thaliana]

Seq. No. 9932

Contig ID 17476_2.R1040

5'-most EST LIB3109-005-Q1-K1-C9

Method BLASTX
NCBI GI g2952433
BLAST score 306
E value 6.0e-28
Match length 87
% identity 84

NCBI Description (AF051135) putative ubiquitin activating enzyme E1

[Arabidopsis thaliana]

Seq. No. 9933

Contig ID 17478 1.R1040

5'-most EST LIB3170-026-Q1-K1-C3

Seq. No. 9934

Contig ID 17479_1.R1040 5'-most EST kll701207665.h1

Method BLASTX
NCBI GI g3695388
BLAST score 250
E value 4.0e-21
Match length 58
% identity 72

NCBI Description (AF096371) No definition line found [Arabidopsis thaliana]

Seq. No. 9935

Contig ID 17481_1.R1040 5'-most EST pcp700989792.h1

Method BLASTX
NCBI GI g4101626
BLAST score 626
E value 5.0e-65
Match length 255
% identity 48

NCBI Description (AF005096) desaturase/cytochrome b5 protein [Ricinus

communis]

Seq. No. 9936

Contig ID 17483_1.R1040 5'-most EST pmv700890620.h1

Seq. No. 9937

Contig ID 17486 1.R1040

5'-most EST LIB3049-020-Q1-E1-H9

Method BLASTX
NCBI GI g3184096
BLAST score 147
E value 2.0e-09



Match length 129 % identity 26

NCBI Description (AL023777) hypothetical protein [Schizosaccharomyces pombe]

Seq. No. 9938

Contig ID 17488 1.R1040

5'-most EST LIB3170-027-Q2-K1-F3

Seq. No. 9939

Contig ID 17492 1.R1040

5'-most EST LIB3049-020-Q1-E1-F7

Method BLASTX
NCBI GI g3420047
BLAST score 244
E value 9.0e-21
Match length 106
% identity 56

NCBI Description (AC004680) putative protein kinase [Arabidopsis thaliana]

Seq. No. 9940

Contig ID 17494 1.R1040

5'-most EST LIB3170-028-Q1-K1-F3

Seq. No. 9941

Contig ID 17498 1.R1040

5'-most EST LIB3049-020-Q1-E1-F9

Seq. No. 9942

Contig ID 17499 1.R1040

5'-most EST LIB3049-020-Q1-E1-G10

Seq. No. 9943

Contig ID 17500_1.R1040

5'-most EST LIB3049-020-Q1-E1-G11

Method BLASTX
NCBI GI g3269286
BLAST score 403
E value 5.0e-39
Match length 132
% identity 66

NCBI Description (AL030978) putative protein [Arabidopsis thaliana]

Seq. No. 9944

Contig ID 17503 1.R1040

5'-most EST LIB $30\overline{4}$ 9-020-Q1-E1-G3

Seq. No. 9945

Contig ID 17508_1.R1040 5'-most EST jex700906877.h1

Method BLASTN
NCBI GI g1370171
BLAST score 317
E value 1.0e-178
Match length 694
% identity 88

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB1X



Contig ID 17509_1.R1040 5'-most EST rlr700898646.h1

Seq. No. 9947

Contig ID 17509_2.R1040

5'-most EST LIB3049-009-Q1-E1-B8

Seq. No. 9948

Contig ID 17512_1.R1040

5'-most EST LIB3049-010-Q1-E1-B11

Seq. No. 9949

Contig ID 17512 2.R1040

5'-most EST LIB3049-033-Q1-E1-B6

Method BLASTN
NCBI GI g3702736
BLAST score 37
E value 2.0e-11

E value 2.0 Match length 77 % identity 87

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MRI1, complete sequence [Arabidopsis thaliana]

Seq. No. 9950

Contig ID 17515_1.R1040

5'-most EST LIB3049-020-Q1-E1-E7

Method BLASTX
NCBI GI g2914703
BLAST score 146
E value 3.0e-09
Match length 27
% identity 100

NCBI Description (AC003974) unknown protein [Arabidopsis thaliana]

Seq. No. 9951

Contig ID 17517_1.R1040

5'-most EST LIB3170-025-Q1-J1-D3

Seq. No. 9952

Contig ID 17519 1.R1040

5'-most EST jC-gmf102220082f03d1

Method BLASTX
NCBI GI g3915039
BLAST score 371
E value 2.0e-35
Match length 124
% identity 56

NCBI Description SUGAR CARRIER PROTEIN C >gi_169718 (L08196) sugar carrier

protein [Ricinus communis]

Seq. No. 9953

Contig ID 17530 1.R1040

5'-most EST uC-gmrominsoy310d04b1

Method BLASTX
NCBI GI g4454464
BLAST score 242



E value 2.0e-20 Match length 123 % identity 46

(AC006234) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No.

9954

Contig ID

17530 2.R1040

5'-most EST

LIB3049-020-Q1-E1-D2

Seq. No.

9955

Contig ID

17531 1.R1040

5'-most EST

LIB3072-047-Q1-K1-B6

Seq. No.

9956

Contig ID 5'-most EST 17533 1.R1040 sat701008660.h1

Method NCBI GI BLAST score BLASTX g2494304 439

E value Match length % identity

3.0e-43178 52

NCBI Description

TRANSLATION INITIATION FACTOR EIF-2B ALPHA SUBUNIT (EIF-2B

GDP-GTP EXCHANGE FACTOR) >gi_2144003_pir__I59376

translation initiation factor eIF-2B alpha-subunit - rat >gi 623033 (U05821) translation initiation factor eIF-2B alpha-subunit [Rattus norvegicus] >gi 1050328 (L41679)

GTP-exchange protein [Rattus norvegicus]

>gi_1096885_prf__2112359A initiation factor eIF-2B [Rattus

norvegicus]

Seq. No.

9957

Contig ID 5'-most EST 17534 1.R1040 fC-qmf1700906023a1

Method BLASTX

NCBI GI BLAST score g4140326 641

E value Match length 2.0e-66

% identity

314 41

NCBI Description

(ALO31282) dJ283E3.6.1 (PUTATIVE novel protein similar to many (archae) bacterial, worm and yeast hypothetical

proteins) [Homo sapiens]

Seq. No.

9958

Contig ID

17535 1.R1040

5'-most EST

uC-gmrominsoy133d06b1

Method NCBI GI BLASTX g462253

BLAST score

165

E value Match length 3.0e-11 151

% identity

27

NCBI Description

TRANSCRIPTIONAL REGULATOR PROTEIN HCNGP

>gi_284975_pir__S26660 HCNGP protein - mouse

>gi_57912_emb_CAA48198_ (X68061) HCNGP [Mus musculus]



Contig ID 17548_1.R1040 5'-most EST uaw700665763.h1

Seq. No. 9960

Contig ID 17557_1.R1040

5'-most EST LIB3109-014-Q1-K1-G8

Seq. No. 9961

Contig ID 17562_1.R1040

5'-most EST LIB3170-026-Q1-K1-F2

Method BLASTN
NCBI GI g3413472
BLAST score 142
E value 5.0e-74
Match length 308
% identity 87

NCBI Description Glycine max mRNA for tyrosine phosphatase

Seq. No. 9962

Contig ID 17563 1.R1040

5'-most EST LIB3049-020-Q1-E1-A1

Method BLASTX
NCBI GI g3510253
BLAST score 373
E value 1.0e-35
Match length 114
% identity 67

NCBI Description (AC005310) hypothetical protein [Arabidopsis thaliana]

Seq. No. 9963

Contig ID 17565_1.R1040 5'-most EST leu701146507.h1

Method BLASTN
NCBI GI g1370177
BLAST score 434
E value 0.0e+00
Match length 796
% identity 92

NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB5A

Seq. No. 9964

Contig ID 17567 1.R1040

5'-most EST LIB3170-027-Q2-K1-G2

Seq. No. 9965

Contig ID 17568 1.R1040

5'-most EST LIB3049-020-Q1-E1-A3

Seq. No. 9966

Contig ID 17570 1.R1040

5'-most EST LIB3107-029-Q1-K1-G9

Seq. No. 9967

Contig ID 17573_1.R1040

5'-most EST LIB3170-027-Q2-K1-B2

Method BLASTX



NCBI GI g731667 BLAST score 153 E value 3.0e-10 Match length 87 % identity 39

NCBI Description ATP-DEPENDENT RRNA HELICASE RRP3 >gi_626623_pir__S46713

hypothetical protein YHR065c - yeast (Saccharomyces cerevisiae) >gi_487961 (U00061) Yhr065cp [Saccharomyces

cerevisiae]

Seq. No. 9968

Contig ID 17575_1.R1040

5'-most EST LIB3049-019-Q1-E1-G1

Method BLASTX
NCBI GI g4580523
BLAST score 185
E value 4.0e-28
Match length 338
% identity 41

NCBI Description (AF036305) scarecrow-like 8 [Arabidopsis thaliana]

Seq. No. 9969

Contig ID 17576_1.R1040

5'-most EST LIB3170-026-Q1-K1-D2

Seq. No. 9970

Contig ID 17578_1.R1040

5'-most EST LIB3049-019-Q1-E1-G12

Seq. No. 9971

Contig ID 17579 1.R1040 5'-most EST kll701214122.h1

Method BLASTX
NCBI GI g3876865
BLAST score 255
E value 2.0e-21
Match length 187
% identity 36

NCBI Description (Z81534) predicted using Genefinder; cDNA EST EMBL: C08177

comes from this gene; cDNA EST EMBL:C09822 comes from this gene; cDNA EST yk359h8.5 comes from this gene; cDNA EST yk374f11.5 comes from this gene; cDNA EST yk359h8.3 co

Seq. No. 9972

Contig ID 17580_1.R1040

5'-most EST LIB3109-055-Q1-K1-C11

Method BLASTX
NCBI GI g2924363
BLAST score 417
E value 1.0e-40
Match length 88
% identity 91

NCBI Description (AJ224683) zeta-carotene desaturase [Narcissus

pseudonarcissus]

Seq. No. 9973

Contig ID 17580_2.R1040



ssr700560212.hl 5'-most EST Method BLASTX g1053093 NCBI GI 189 BLAST score E value 4.0e-14 69 Match length 59 % identity (U38550) zeta-carotene desaturase precursor [Arabidopsis NCBI Description thaliana] Seq. No. 9974 17581 1.R1040 Contig ID 5'-most EST LIB3049-019-Q1-E1-G5 Seq. No. 9975 17582 1.R1040 Contig ID 5'-most EST bth700847081.h1 Method BLASTX NCBI GI g2160150 BLAST score 346 E value 2.0e-32 Match length 163 48 % identity (ACO00375) EST gb T43829 comes from this gene. [Arabidopsis NCBI Description thaliana] 9976 Seq. No. 17582 2.R1040 Contig ID LIB3109-014-Q1-K1-A6 5'-most EST 9977 Seq. No. Contig ID 17583 1.R1040 5'-most EST LIB3049-019-Q1-E1-G7 Method BLASTX g3075399 NCBI GI BLAST score 157 E value 1.0e-10 97 Match length % identity 38 (AC004484) SF16-like protein [Arabidopsis thaliana] NCBI Description 9978 Seq. No. 17584 1.R1040 Contig ID 5'-most EST LIB3049-019-Q1-E1-G8 Method BLASTX g192185 NCBI GI BLAST score 429 E value 4.0e-42 Match length 155 % identity 52 NCBI Description (M75122) acid beta-galactosidase [Mus musculus]

9979

Contig ID 17590_1.R1040

5'-most EST LIB3087-007-Q1-K1-C3 Method BLASTX

Method BLASTX NCBI GI g1711507

Seq. No.



BLAST score 338 E value 1.0e-31 Match length 106 % identity 62

NCBI Description SIGNAL RECOGNITION PARTICLE 19 KD PROTEIN (SRP19)

>gi_624221 (U19030) signal recognition particle 19 kDa

protein subunit SRP19 [Oryza sativa]

Seq. No. 9980

Contig ID 17597_1.R1040 5'-most EST sat701005120.h1

Seq. No. 9981

Contig ID 17603_1.R1040 5'-most EST kl1701202580.h1

Seq. No. 9982

Contig ID 17606 1.R1040 5'-most EST jex700904276.h1

Method BLASTX
NCBI GI g3355477
BLAST score 357
E value 2.0e-33
Match length 96
% identity 45

NCBI Description (AC004218) putative P-glycoprotein, pgp1 [Arabidopsis

thaliana]

Seq. No. 9983

Contig ID 17607_1.R1040 5'-most EST jsh701066283.h1

Method BLASTN
NCBI GI g2564049
BLAST score 33
E value 5.0e-09
Match length 109
% identity 83

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MLE2, complete sequence [Arabidopsis thaliana]

Seq. No. 9984

Contig ID 17609_1.R1040

5'-most EST LIB3049-019-Q1-E1-C8

Method BLASTX
NCBI GI g4539359
BLAST score 239
E value 4.0e-20
Match length 81
% identity 53

NCBI Description (AL049525) putative protein [Arabidopsis thaliana]

Seq. No. 9985

Contig ID 17610_1.R1040 5'-most EST uC-gmropic088g05b1

Method BLASTN
NCBI GI g2924733
BLAST score 55



E value 8.0e-22 Match length 172 % identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MUF9, complete sequence [Arabidopsis thaliana]

Seq. No. 9986

Contig ID 17612 1.R1040

5'-most EST LIB3139-031-P1-N1-F7

Method BLASTX
NCBI GI g3367522
BLAST score 1074
E value 1.0e-117
Match length 322
% identity 64

NCBI Description (AC004392) EST gb T04691 comes from this gene. [Arabidopsis

thaliana)

Seq. No.

Contig ID 17619_1.R1040

5'-most EST LIB3065-027-Q1-N1-B1

9987

Method BLASTX
NCBI GI g1050918
BLAST score 302
E value 4.0e-27
Match length 92
% identity 60

NCBI Description (X92648) lipid transfer protein [Helianthus annuus]

Seq. No. 9988

Contig ID 17619 2.R1040

5'-most EST LIB3170-014-Q1-K1-G11

Method BLASTX
NCBI GI g3914136
BLAST score 284
E value 5.0e-25
Match length 100
% identity 56

NCBI Description NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (LTP)

>gi 2632171 emb CAA05771 (AJ002958) lipid transfer protein

[Cicer arietinum]

Seq. No. 9989

Contig ID 17619_3.R1040 5'-most EST uaw700665883.h1

Method BLASTX
NCBI GI g2407271
BLAST score 168
E value 6.0e-12
Match length 78
% identity 49

NCBI Description (AF017358) lipid transfer protein [Oryza sativa]

Seq. No. 9990

Contig ID 17621 1.R1040

5'-most EST LIB3049-019-Q1-E1-C2



Contig ID 17622 1.R1040

Method BLASTX
NCBI GI g3875304
BLAST score 556
E value 5.0e-57
Match length 164
% identity 60

NCBI Description (Z74030) predicted using Genefinder; cDNA EST EMBL:C07609 comes from this gene; cDNA EST EMBL:C09023 comes from this gene; cDNA EST yk505e9.3 comes from this gene; cDNA EST

yk489h9.3 comes from this gene; cDNA EST yk489h9.5 com

Seq. No. 9992

Contig ID 17632 1.R1040

5'-most EST LIB3049-019-Q1-E1-B8

Seq. No. 9993

Contig ID 17633 1.R1040

5'-most EST LIB3170-028-Q1-K1-D2

Seq. No. 9994

Contig ID 17651 1.R1040

5'-most EST LIB3049-019-Q1-E1-A3

Seq. No. 9995

Contig ID 17654 1.R1040

5'-most EST LIB3049-018-Q1-E1-F3

Seq. No. 9996

Contig ID 17655 1.R1040

5'-most EST LIB3170-027-Q2-K1-E1

Seq. No. 9997

Contig ID 17663 1.R1040

5'-most EST LIB3170-047-Q1-J1-C2

Seq. No. 9998

Contig ID 17665 1.R1040

5'-most EST LIB $30\overline{4}9-018-Q1-E1-G7$

Seq. No. 9999

Contig ID 17666_1.R1040 5'-most EST pcp700990149.h1

Seq. No. 10000

Contig ID 17672 1.R1040

5'-most EST LIB3049-018-Q1-E1-E12

Seq. No. 10001

Contig ID 17673_1.R1040

5'-most EST LIB3049-018-Q1-E1-E2

Seq. No. 10002

Contig ID 17673_2.R1040 5'-most EST rca700998058.h1



```
BLASTX
Method
                  q2623295
NCBI GI
                  247
BLAST score
                   3.0e-22
E value
                   107
Match length
% identity
                   59
NCBI Description (AC002409) hypothetical protein [Arabidopsis thaliana]
Seq. No.
                   10003
                   17675 1.R1040
Contig ID
                   leu701147592.h1
5'-most EST
                   10004
Seq. No.
                   17675 2.R1040
Contig ID
                   sat701007648.hl
5'-most EST
Seq. No.
                   10005
                   17676 1.R1040
Contig ID
                   LIB3049-018-Q1-E1-E5
5'-most EST
                   BLASTX
Method
                   q2245086
NCBI GI
                   581
BLAST score
                   8.0e-60
E value
                   144
Match length
% identity
                   (Z97343) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10006
Seq. No.
                   17688 1.R1040
Contig ID
                   kl1701203922.h2
5'-most EST
                   10007
Seq. No.
                   17689 1.R1040
Contig ID
                   LIB3170-025-Q1-K1-F1
 5'-most EST
                   BLASTX
Method
                   g2104446
NCBI GI
BLAST score
                   213
                   1.0e-16
 E value
                   271
 Match length
                   24
 % identity
                  (Z95396) WD-repeat protein [Schizosaccharomyces pombe]
 NCBI Description
                   10008
 Seq. No.
                   17694 1.R1040
 Contig ID
                   fua701042211.h1
 5'-most EST
                   BLASTN
 Method
                    g4519187
 NCBI GI
                    37
 BLAST score
                    4.0e-11
 E value
                    106
 Match length
                    89
 % identity
 NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone:
                    K1G2, complete sequence
```

17704 1.R1040 Contig ID

LIB3049-018-Q1-E1-A12 5'-most EST



Method BLASTX
NCBI GI g2492511
BLAST score 580
E value 7.0e-60
Match length 182
% identity 63

NCBI Description CELL DIVISION PROTEIN FTSH HOMOLOG 2

>gi_1652282_dbj_BAA17205_ (D90904) cell division protein

FtsH [Synechocystis sp.]

Seq. No. 10010

Contig ID 17719_1.R1040

5'-most EST LIB3049-018-Q1-E1-A11

Seq. No. 10011

Contig ID 17720_1.R1040

5'-most EST LIB3170-025-Q1-K1-E1

Seq. No. 10012

Contig ID 17721_1.R1040 5'-most EST eep700864833.h1

Seq. No. 10013

Contig ID 17723_1.R1040

5'-most EST LIB3170-021-Q1-K1-H11

Method BLASTX
NCBI GI g4531442
BLAST score 223
E value 3.0e-18
Match length 58
% identity 78

NCBI Description (AC006224) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10014

Contig ID 17725 1.R1040

5'-most EST LIB3170-022-Q1-K1-H11

Seq. No. 10015

Contig ID 17736_1.R1040

5'-most EST LIB3170-028-Q1-K1-B1

Seq. No. 10016

Contig ID 17738_1.R1040

5'-most EST LIB3049-017-Q1-E1-G12

Seq. No. 10017

Contig ID 17740 1.R1040

5'-most EST LIB3170-025-Q1-K1-D1

Seq. No. 10018

Contig ID 17744_1.R1040

5'-most EST LIB3049-017-Q1-E1-F1

Seq. No. 10019

Contig ID 17755_1.R1040 5'-most EST seb700649613.h1

Method BLASTX



```
g4539306
NCBI GI
BLAST score
                  147
                   3.0e-09
E value
                  120
Match length
% identity
                  (AL049480) pumilio-like protein [Arabidopsis thaliana]
NCBI Description
                   10020
Seq. No.
                   17761 1.R1040
Contig ID
                   g5510363
5'-most EST
                   BLASTX
Method
                   q1653142
NCBI GI
                   213
BLAST score
                   1.0e-16
E value
                   135
Match length
                   39
% identity
                  (D90911) hypothetical protein [Synechocystis sp.]
NCBI Description
                   10021
Seq. No.
                   17762 1.R1040
Contig ID
                   LIB3049-017-Q1-E1-D11
5'-most EST
                   10022
Seq. No.
                   17765 1.R1040
Contig ID
                   jC-qmro02910020d12a1
5'-most EST
Seq. No.
                   10023
                   17765 2.R1040
Contig ID
                   uC-qmrominsoy025d02b1
5'-most EST
                   10024
Seq. No.
                   17770 1.R1040
Contig ID
                   xpa700793759.h1
 5'-most EST
                   BLASTX
Method
                   q2897875
NCBI GI
                   228
 BLAST score
                   8.0e-19
 E value
                   79
Match length
                   52
 % identity
                   (U90274) histone acetyltransferase HAT B [Zea mays]
 NCBI Description
                   10025
 Seq. No.
                   17771 1.R1040
 Contig ID
                   epx701110128.h1
 5'-most EST
                    10026
 Seq. No.
                    17772 1.R1040
 Contig ID
                   LIB3170-022-Q1-K1-H3
 5'-most EST
                    10027
 Seq. No.
                    17774 1.R1040
```

Contig ID

LIB3170-021-Q1-J1-H9 5'-most EST

10028 Seq. No.

17777 1.R1040 Contig ID

5'-most EST LIB3049-017-Q1-E1-A9

BLASTX Method



```
g2252840
NCBI GI
BLAST score
                  248
E value
                  3.0e-21
Match length
                  105
% identity
                   45
                   (AF013293) contains regions of similarity to Haemophilus
NCBI Description
                  influenzae permease (SP:P38767) [Arabidopsis thaliana]
                   10029
Seq. No.
                   17781 1.R1040
Contig ID
5'-most EST
                  ncj700982438.hl
                   10030
Seq. No.
Contig ID
                   17783 1.R1040
5'-most EST
                   q5688018
Method
                   BLASTX
NCBI GI
                   q3549665
BLAST score
                   414
                   4.0e-40
E value
                   373
Match length
                   33
% identity
                  (AL031394) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10031
Seq. No.
                   17785 1.R1040
Contig ID
5'-most EST
                   kl1701203177.hl
Method
                   BLASTX
                   g3928150
NCBI GI
BLAST score
                   628
E value
                   1.0e-65
Match length
                   161
                   75
% identity
                  (AJ131049) hypothetical protein [Cicer arietinum]
NCBI Description
                   10032
Seq. No.
                   17786 1.R1040
Contig ID
5'-most EST
                   uC-gmropic056f06b1
Method
                   BLASTX
NCBI GI
                   g2243118
BLAST score
                   1170
                   1.0e-128
E value
Match length
                   316
% identity
                   68
                   (Y10984) glutathione synthetase [Brassica juncea]
NCBI Description
                   10033
Seq. No.
                   17786 2.R1040
Contig ID
                   LIB3093-018-Q1-K2-H4
5'-most EST
                   BLASTX
Method
                   g2243118
NCBI GI
```

Method BLASTX
NCBI GI g2243118
BLAST score 193
E value 1.0e-22
Match length 87
% identity 66

NCBI Description (Y10984) glutathione synthetase [Brassica juncea]

Seq. No. 10034



17787 1.R1040 Contig ID LIB3049-017-Q1-E1-B9 5'-most EST Method BLASTX NCBI GI g4432840 BLAST score 191 2.0e-14 E value Match length 67 51 % identity NCBI Description

(AC006283) unknown protein [Arabidopsis thaliana]

10035 Seq. No.

17789 1.R1040 Contig ID

LIB3049-017-Q1-E1-C10 5'-most EST

10036 Seq. No.

Contig ID 17796 1.R1040 5'-most EST zhf700960425.h1

10037 Seq. No.

17796 2.R1040 Contig ID 5'-most EST leu701146313.h1

10038 Seq. No.

17798 1.R1040 Contig ID

5'-most EST LIB3093-025-Q1-K1-H8

Seq. No. 10039

17798 2.R1040 Contig ID

5'-most EST LIB3170-022-Q1-K1-H4

10040 Seq. No.

Contig ID 17801 1.R1040 ncj700983102.hl 5'-most EST

10041 Seq. No.

17804 1.R1040 Contig ID

LIB3049-017-Q1-E1-A12 5'-most EST

BLASTX Method g3242785 NCBI GI 552 BLAST score 2.0e-56 E value Match length 129 % identity 80

(AF055355) respiratory burst oxidase protein C [Arabidopsis NCBI Description

thaliana]

10042 Seq. No.

17806 1.R1040 Contig ID

jC-gmf102220072g04a1 5'-most EST

10043 Seq. No.

Contig ID 17807 1.R1040

LIB3049-017-Q1-E1-A4 5'-most EST

BLASTX Method g2245141 NCBI GI BLAST score 452 5.0e-45 E value

1765



Match length 105 % identity 77 (Z97344) acetylornithine deacetylase [Arabidopsis thaliana] NCBI Description Seq. No. 10044 Contig ID 17815 1.R1040 pxt700944221.h1 5'-most EST BLASTX Method q2673920 NCBI GI 232 BLAST score E value 5.0e-19 97 Match length % identity 49 (AC002561) similar to Drosophila couch potato protein NCBI Description [Arabidopsis thaliana] Seq. No. 10045 Contig ID 17815 2.R1040 LIB3049-016-Q1-E1-F9 5'-most EST BLASTX Method q2673920 NCBI GI BLAST score 159 1.0e-10 E value 67 Match length % identity 49 (ACO02561) similar to Drosophila couch potato protein NCBI Description [Arabidopsis thaliana] 10046 Seq. No. Contig ID 17818 1.R1040 LIB3170-050-Q1-K1-A7 5'-most EST BLASTX Method g3858937 NCBI GI 649 BLAST score 7.0e-68 E value 216 Match length % identity (AL021636) putative protein [Arabidopsis thaliana] NCBI Description 10047 Seq. No. 17819 1.R1040 Contig ID LIB3051-032-Q1-K1-E6 5'-most EST 10048 Seq. No. 17825_1.R1040 Contig ID 5'-most EST LIB3049-016-Q1-E1-G8 10049 Seq. No. 17827 1.R1040 Contig ID rca700999233.h1 5'-most EST BLASTN Method g3176071 NCBI GI

Method BLASTN
NCBI GI g317607
BLAST score 555
E value 0.0e+00
Match length 1085
% identity 90

NCBI Description Medicago sativa mRNA for protein phosphatase 1, beta



subunit

10050 Seq. No. Contig ID 17827 2.R1040 dpv701101871.h1 5'-most EST BLASTN Method g3176071 NCBI GI 114 BLAST score 4.0e-57 E value 242 Match length % identity 87 NCBI Description subunit

Medicago sativa mRNA for protein phosphatase 1, beta

10051 Seq. No.

Contig ID 17827 3.R1040

LIB3106-074-Q1-K1-H3 5'-most EST

BLASTN Method g3176071 NCBI GI 165 BLAST score E value 1.0e-87 332 Match length % identity 92

Medicago sativa mRNA for protein phosphatase 1, beta NCBI Description

subunit

10052 Seq. No.

17827 4.R1040 Contig ID $k1170\overline{1}214536.h1$ 5'-most EST

BLASTX Method g3176074 NCBI GI 480 BLAST score 2.0e-48 E value 138 Match length 70 % identity

(AJ002486) protein phosphatase 1, catalytic gsmms subunit NCBI Description

[Medicago sativa]

10053 Seq. No.

17829 1.R1040 Contig ID

LIB3138-034-Q1-N1-E3 5'-most EST

BLASTX Method g3860250 NCBI GI 277 BLAST score 2.0e-24 E value Match length 83 59 % identity

(AC005824) putative chloroplast prephenate dehydratase NCBI Description

[Arabidopsis thaliana]

Seq. No. 10054

17829 2.R1040 Contig ID

jC-qmle01810019d08a2 5'-most EST

BLASTX Method g3860250 NCBI GI BLAST score 271 E value 1.0e-23

1767



Match length 90 % identity 53

NCBI Description (AC005824) putative chloroplast prephenate dehydratase

[Arabidopsis thaliana]

Seq. No. 10055

Contig ID 17830_1.R1040

5'-most EST LIB3049-016-Q1-E1-E10

Method BLASTX
NCBI GI g1621438
BLAST score 125
E value 3.0e-09
Match length 138
% identity 30

NCBI Description (U71603) mitochondrial transport protein amc-1 [Emericella

nidulans]

Seq. No. 10056

Contig ID 17832_1.R1040

5'-most EST LIB3049-016-Q1-E1-E12

Method BLASTX
NCBI GI 94454480
BLAST score 279
E value 2.0e-24
Match length 90
% identity 61

NCBI Description (AC006234) putative (1-4)-beta-mannan endohydrolase

[Arabidopsis thaliana]

Seq. No. 10057

Contig ID 17839_1.R1040 5'-most EST epx701105337.h1

Seq. No. 10058

Contig ID 17841 2.R1040

5'-most EST jC-gmle01810020a05d1

Seq. No. 10059

Contig ID 17841 3.R1040

5'-most EST LIB3049-016-Q1-E1-F10

Method BLASTX
NCBI GI g1174850
BLAST score 743
E value 9.0e-79
Match length 149
% identity 92

NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-21 KD 2 (UBIQUITIN-PROTEIN

LIGASE 5) (UBIQUITIN CARRIER PROTEIN 5)

Seq. No. 10060

Contig ID 17841_4.R1040

5'-most EST LIB3109-047-Q1-K1-D8

Method BLASTX
NCBI GI g1174850
BLAST score 753
E value 7.0e-80
Match length 149



% identity 93
NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-21 KD 2 (UBIQUITIN-PROTEIN
LIGASE 5) (UBIQUITIN CARRIER PROTEIN 5)

Seq. No. 10061

Contig ID 17841_5.R1040 5'-most EST sat701006343.h1

Method BLASTX
NCBI GI g431270
BLAST score 390
E value 1.0e-37
Match length 80
% identity 89

NCBI Description (L19356) ubiquitin conjugating enzyme [Arabidopsis

thaliana]

Seq. No. 10062

Contig ID 17841_6.R1040 5'-most EST zhf700954546.h1

Method BLASTX
NCBI GI g136644
BLAST score 202
E value 5.0e-16
Match length 42
% identity 88

NCBI Description UBIQUITIN-CONJUGATING ENZYME E2-23 KD (UBIQUITIN-PROTEIN

LIGASE) (UBIQUITIN CARRIER PROTEIN) >gi_100765_pir__A34506

23K ubiquitin carrier protein E2 - wheat >gi_170782 (M28059) ubiquitin carrier protein [Triticum vulgare]

Seq. No. 10063

Contig ID 17841_7.R1040

5'-most EST uC-qmronoir023h05b1

Method BLASTN
NCBI GI g431269
BLAST score 46
E value 8.0e-17
Match length 105
% identity 92

NCBI Description Arabidopsis thaliana ubiquitin conjugating enzyme (UBC5)

gene, complete cds

Seq. No. 10064

Contig ID 17843_1.R1040 5'-most EST wrg700786292.h2

Method BLASTX
NCBI GI g4406815
BLAST score 731
E value 2.0e-77
Match length 209
% identity 67

NCBI Description (AC006201) putative peptide methionine sulfoxide reductase

[Arabidopsis thaliana]

Seq. No. 10065

Contig ID 17843 2.R1040

5'-most EST LIB3170-021-Q1-K1-H2



```
BLASTX
Method
                  g4406815
NCBI GI
                  275
BLAST score
                  3.0e-24
E value
Match length
                  81
                   64
% identity
                   (AC006201) putative peptide methionine sulfoxide reductase
NCBI Description
                   [Arabidopsis thaliana]
                  10066
Seq. No.
                  17845 2.R1040
Contig ID
                  pxt700946337.h1
5'-most EST
                   BLASTX
Method
                   q3288888
NCBI GI
BLAST score
                   145
E value
                   4.0e-09
                   116
Match length
                   28
% identity
NCBI Description (AC005253) R26445_1 [Homo sapiens]
Seq. No.
                   10067
                   17849 1.R1040
Contig ID
                   trc700565235.h1
5'-most EST
                   BLASTX
Method
                   g4559356
NCBI GI
BLAST score
                   163
                   6.0e-11
E value
Match length
                   146
% identity
                   10
                  (AC006585) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10068
Seq. No.
                   17855 1.R1040
Contig ID
                   jC-gmst02400076f04a1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g4218121
                   148
BLAST score
E value
                   3.0e-09
Match length
                   43
                   56
% identity
NCBI Description (AL035353) putative protein [Arabidopsis thaliana]
                   10069
Seq. No.
                   17855 2.R1040
Contig ID
5'-most EST
                   LIB3049-016-Q1-E1-D10
Method
                   BLASTX
NCBI GI
                   g4218121
BLAST score
                   413
                   3.0e-40
E value
Match length
                   152
% identity
NCBI Description (AL035353) putative protein [Arabidopsis thaliana]
```

Seq. No. 10070

Contig ID 17856 1.R1040

5'-most EST LIB3107-035-Q1-K1-E9

Method BLASTX



g4206210 NCBI GI BLAST score 336 2.0e-31 E value Match length 128 % identity 54

(AF071527) putative calcium channel [Arabidopsis thaliana] NCBI Description

>qi 4263043 gb AAD15312 (AC005142) putative calcium

channel [Arabidopsis thaliana]

10071 Seq. No.

17860 1.R1040 Contig ID

LIB3049-016-Q1-E1-D5 5'-most EST

10072 Seq. No.

Contig ID 17861 1.R1040

LIB3049-016-Q1-E1-D6 5'-most EST

10073 Seq. No.

17863 1.R1040 Contig ID vzy700755952.h1 5'-most EST

Method BLASTX q465820 NCBI GI BLAST score 587 1.0e-60 E value Match length 161 % identity 63

HYPOTHETICAL 18.5 KD PROTEIN C40H1.6 IN CHROMOSOME III NCBI Description

>gi_280536_pir__S28301 hypothetical protein C40H1.6 -Caenorhabditis elegans >gi_3874819_emb_CAA79557 (Z19154)

C40H1.6 [Caenorhabditis elegans]

10074 Seq. No.

17868 1.R1040 Contig ID leu701156625.h1 5'-most EST

BLASTX Method g3860273 NCBI GI 172 BLAST score 2.0e-12 E value 95 Match length % identity 45

(AC005824) hypothetical protein [Arabidopsis thaliana] NCBI Description >gi_4314398_gb_AAD15608_ (AC006232) hypothetical protein

[Arabidopsis thaliana]

10075 Seq. No.

Contig ID 17869 1.R1040

5'-most EST LIB3170-023-Q1-K1-H5

10076 Seq. No.

Contig ID 17881_1.R1040 5'-most EST leu701144861.h1

Method BLASTX q1255954 NCBI GI 490 BLAST score E value 3.0e-49 Match length 116 77 % identity



```
(Z70677) thioredoxin [Ricinus communis]
NCBI Description
                  10077
Seq. No.
Contiq ID
                  17881 2.R1040
                  jC-qmst02400004a08d1
5'-most EST
                  BLASTX
Method
                  q586099
NCBI GI
                  224
BLAST score
                  4.0e-18
E value
Match length
                  54
% identity
                  78
                  THIOREDOXIN H-TYPE 2 (TRX-H2) >gi_486690_pir__S34812
NCBI Description
                  thioredoxin h2 - common tobacco >gi_297519_emb_CAA77847
                   (Z11803) THIOREDOXIN [Nicotiana tabacum]
                  >gi 447151 prf 1913431A thioredoxin [Nicotiana tabacum]
Seq. No.
                  10078
                  17882 1.R1040
Contig ID
                  hyd700727023.hl
5'-most EST
Method
                  BLASTX
                  g1200205
NCBI GI
                   347
BLAST score
                   3.0e-32
E value
                   124
Match length
                   55
% identity
                  (X95753) DAG [Antirrhinum majus]
NCBI Description
Seq. No.
                   10079
                   17886 1.R1040
Contig ID
                   LIB3049-015-Q1-E1-H6
5'-most EST
                   10080
Seq. No.
                   17888 1.R1040
Contig ID
                   uC-gmropic022f04b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g2632252
                   761
BLAST score
                   7.0e-81
E value
                   199
Match length
                   70
% identity
NCBI Description (Y12464) serine/threonine kinase [Sorghum bicolor]
                   10081
Seq. No.
                   17888 2.R1040
Contig ID
5'-most EST
                   kmv700740252.h1
                   BLASTX
Method
                   q3885328
NCBI GI
BLAST score
                   1525
                   1.0e-170
E value
Match length
                   338
% identity
                   83
                  (AC005623) putative serine/threonine protein kinase
NCBI Description
```

Seq. No. 10082

Contig ID 17888_3.R1040

5'-most EST LIB3138-127-Q1-N1-C4

[Arabidopsis thaliana]



```
10083
Seq. No.
Contig ID
                  17888 4.R1040
5'-most EST
                  jC-gmle01810044e03a1
Seq. No.
                  10084
                  17888 6.R1040
Contig ID
                  jC-gmle01810075e03d1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3885328
BLAST score
                  210
                  1.0e-16
E value
Match length
                   43
                   95
% identity
                   (AC005623) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   10085
                   17888 8.R1040
Contig ID
                   zzp700831282.h1
5'-most EST
Method
                   BLASTX
                   q3885328
NCBI GI
                   207
BLAST score
                   2.0e-16
E value
Match length
                   43
% identity
                   93
                   (AC005623) putative serine/threonine protein kinase
NCBI Description
                   [Arabidopsis thaliana]
                   10086
Seq. No.
                   17889 1.R1040
Contig ID
                   jC-gm\overline{f}102220138c12a1
5'-most EST
                   BLASTX
Method
                   g4220474
NCBI GI
BLAST score
                   266
                   5.0e-23
E value
                   92
Match length
                   68
% identity
                   (AC006069) putative myosin heavy chain [Arabidopsis
NCBI Description
                   thaliana]
                   10087
Seq. No.
                   17903 1.R1040
Contig ID
                   fua701043040.h1
5'-most EST
Seq. No.
                   10088
                   17903 2.R1040
Contig ID
                   LIB3093-016-Q1-K1-H8
5'-most EST
                   10089
Seq. No.
Contig ID
                   17907 1.R1040
```

5'-most EST LIB3049-015-Q1-E1-G5

BLASTX Method NCBI GI g2148045 BLAST score 165 2.0e-11 E value Match length 120



% identity
NCBI Description

self-incompatibility protein S3 precursor - Papaver rhoeas >gi_1107841_emb_CAA60578_ (X87100) S3 self-incompatibility protein [Papaver rhoeas] >gi_1107843_emb_CAA60579_ (X87101)

S3 self-incompatibility protein [Papaver rhoeas]

Seq. No. 10090

Contig ID 17912_1.R1040 5'-most EST jsh701069991.h2

Method BLASTX
NCBI GI g3858935
BLAST score 652
E value 4.0e-68
Match length 212
% identity 53

NCBI Description (AL021636) synaptobrevin-like protein [Arabidopsis

thaliana] >gi_4103357 (AF025332) vesicle-associated membrane protein 7C; synaptobrevin 7C [Arabidopsis

thaliana]

Seq. No. 10091

Contig ID 17912_2.R1040 5'-most EST jsh701069609.h1

Method BLASTX
NCBI GI g3858935
BLAST score 182
E value 1.0e-13
Match length 55
% identity 58

NCBI Description (AL021636) synaptobrevin-like protein [Arabidopsis

thaliana] >gi_4103357 (AF025332) vesicle-associated membrane protein 7C; synaptobrevin 7C [Arabidopsis

thaliana]

Seq. No. 10092

Contig ID 17913_1.R1040

5'-most EST LIB3170-055-Q1-K1-F2

Method BLASTX
NCBI GI g2315449
BLAST score 415
E value 3.0e-40
Match length 231
% identity 43

NCBI Description (AF016448) similar to Saccharomyces cerevisiae nuclear

protein SNF7 (SP:P39929)in one region and the chromosome

segregation protein SMC2 (SP:P38989) in another

[Caenorhabditis elegans]

Seq. No. 10093

Contig ID 17913_2.R1040 5'-most EST gsv701053216.h1

Method BLASTX
NCBI GI g3218396
BLAST score 191
E value 2.0e-14
Match length 74
% identity 53

BLAST score

Match length

NCBI Description

% identity

E value

1113

295

74

1.0e-122



```
(AL023860) hypothetical protein [Schizosaccharomyces pombe]
NCBI Description
                  10094
Seq. No.
Contig ID
                  17921 1.R1040
5'-most EST
                  jC-qmf102220092c07d1
                  10095
Seq. No.
Contig ID
                  17923 1.R1040
5'-most EST
                  uC-gmrominsoy186d03b1
                  BLASTX
Method
NCBI GI
                  g2129698
BLAST score
                  780
                  3.0e-83
E value
Match length
                  184
                  80
% identity
                  protein kinase ATN1 (EC 2.7.1.-) - Arabidopsis thaliana
NCBI Description
                  >gi 1054633 emb CAA63387 (X92728) protein kinase
                   [Arabidopsis thaliana]
                  10096
Seq. No.
                  17926 1.R1040
Contig ID
5'-most EST
                  LIB3049-015-Q1-E1-E8
                  10097
Seq. No.
Contig ID
                  17928 1.R1040
5'-most EST
                  pxt700941235.hl
Method
                  BLASTX
NCBI GI
                  q3150410
BLAST score
                   363
E value
                   5.0e-34
Match length
                   236
% identity
                   33
                  (AC004165) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10098
Seq. No.
                   17928 3.R1040
Contig ID
5'-most EST
                  LIB3049-015-Q1-E1-D4
                   10099
Seq. No.
                   17932 1.R1040
Contig ID
5'-most EST
                  LIB3170-022-Q1-K1-G4
Seq. No.
                   10100
Contig ID
                   17933 1.R1040
5'-most EST
                   LIB3049-015-Q1-E1-C12
Seq. No.
                   10101
                   17938 1.R1040
Contig ID
5'-most EST
                   jC-gmst02400050c04a1
                   BLASTX
Method
NCBI GI
                   g3643603
```

1775

(AC005395) unknown protein [Arabidopsis thaliana]



```
10102
Seq. No.
Contig ID
                  17941 1.R1040
                  LIB3049-015-Q1-E1-D1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g2213594
BLAST score
                  307
E value
                  5.0e-28
Match length
                  64
% identity
                  (ACO00348) T7N9.14 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  10103
                  17943 1.R1040
Contig ID
                  epx701107744.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q585973
BLAST score
                  179
E value
                  1.0e-12
                  110
Match length
                   40
% identity
                  FRUCTOKINASE >gi 626018 pir S39997 fructokinase (EC
NCBI Description
                   2.7.1.4) - potato >qi 297015 emb CAA78283 (Z12823)
                   fructokinase [Solanum tuberosum] >gi_1095321_prf__2108342A
                   fructokinase [Solanum tuberosum]
                   10104
Seq. No.
Contig ID
                   17945 1.R1040
                   LIB3049-015-Q1-E1-D2
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4539351
BLAST score
                   369
E value
                   4.0e-35
Match length
                   154
% identity
                   50
NCBI Description (AL035539) putative protein [Arabidopsis thaliana]
Seq. No.
                   10105
                   17946 1.R1040
Contig ID
                   uC-gmflminsoy022c08b1
5'-most EST
Method
                   BLASTX
                   q4432863
NCBI GI
BLAST score
                   179
                   5.0e-13
E value
                   60
Match length
                   53
% identity
                   (AC006300) putative phosphate/phosphoenolpyruvate
NCBI Description
                   translocator protein [Arabidopsis thaliana]
                   10106
Seq. No.
                   17955 1.R1040
Contig ID
                   fde700874633.hl
5'-most EST
```

Seq. No. 10107

17956 1.R1040 Contig ID 5'-most EST $gsv70\overline{1}053913.h1$

Seq. No. 10108



Contig ID 17962_1.R1040 5'-most EST sat701011715.h1

Seq. No. 10109

Contig ID 17962_2.R1040

5'-most EST LIB3049-015-Q1-E1-B2

Method BLASTX
NCBI GI 94210451
BLAST score 150
E value 2.0e-09
Match length 57
% identity 56

NCBI Description (AB016472) ARR2 protein [Arabidopsis thaliana]

Seq. No. 10110

Contig ID 17966_1.R1040 5'-most EST uC-gmropic112d10b1

Method BLASTX
NCBI GI g2245082
BLAST score 156
E value 7.0e-10
Match length 82
% identity 41

NCBI Description (Z97343) SCARECROW homolog [Arabidopsis thaliana]

Seq. No. 10111

Contig ID 17966 2.R1040

5'-most EST LIB3170-023-Q1-K1-G5

Seq. No. 10112

Contig ID 17966_3.R1040 5'-most EST smc700747161.h1

Seq. No. 10113

Contig ID 17969 1.R1040

5'-most EST LIB3049-013-Q1-E1-D7

Method BLASTX
NCBI GI g2252628
BLAST score 183
E value 1.0e-13
Match length 57
% identity 72

NCBI Description (U95973) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10114

Contig ID 17970 1.R1040

5'-most EST uC-gmrominsoy307f04b1

Seq. No. 10115

Contig ID 17971_1.R1040 5'-most EST kl1701210033.h1

Seq. No. 10116

Contig ID 17971_2.R1040

5'-most EST LIB3049-014-Q1-E1-G11

Seq. No. 10117

1777



Contig ID 17976_1.R1040 5'-most EST leu701149308.h1

Method BLASTX
NCBI GI g2829924
BLAST score 790
E value 1.0e-111
Match length 353
% identity 61

NCBI Description (AC002291) Unknown protein [Arabidopsis thaliana]

Seq. No. 10118

Contig ID 17980_1.R1040
5'-most EST g5606496
Method BLASTX
NCBI GI g4176522
BLAST score 198
E value 6.0e-15
Match length 227
% identity 25

NCBI Description (AL035263) hypothetical protein [Schizosaccharomyces pombe]

i

Seq. No. 10119

Contig ID 17980 2.R1040 5'-most EST trc700563873.h1

Seq. No. 10120

Contig ID 17983 1.R1040

5'-most EST jC-gmst02400043d01a1

Seq. No. 10121

Contig ID 17988 1.R1040

5'-most EST LIB3170-023-Q1-K1-G7

Seq. No. 10122

Contig ID 17989_1.R1040

5'-most EST LIB3093-016-Q1-K1-D4

Method BLASTN
NCBI GI 94263540
BLAST score 35
E value 5.0e-10
Match length 78
% identity 96

NCBI Description Arabidopsis thaliana chromosome II BAC T6A13 genomic sequence, complete sequence [Arabidopsis thaliana]

Seq. No. 10123

Contig ID 17990_1.R1040 5'-most EST uC-gmropic006b11b1

Method BLASTX
NCBI GI g2245032
BLAST score 286
E value 2.0e-25
Match length 116
% identity 51

NCBI Description (Z97342) gibberellin oxidase homolog [Arabidopsis thaliana]

Seq. No. 10124



```
17994 1.R1040
Contig ID
                  uC-gmropic070g10b1
5'-most EST
                  BLASTX
Method
                  q2435511
NCBI GI
BLAST score
                  513
                  1.0e-51
E value
                  190
Match length
                  55
% identity
                   (AF024504) contains similarity to prolyl 4-hydroxylase
NCBI Description
                  alpha subunit [Arabidopsis thaliana]
                   10125
Seq. No.
                   17994 2.R1040
Contig ID
                   leu701146423.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2435511
BLAST score
                   452
                   1.0e-46
E value
                   192
Match length
                   54
% identity
                   (AF024504) contains similarity to prolyl 4-hydroxylase
NCBI Description
                   alpha subunit [Arabidopsis thaliana]
                   10126
Seq. No.
                   17997 1.R1040
Contig ID
                   k11701215377.h1
5'-most EST
Method
                   BLASTX
                   q1946355
NCBI GI
                   487
BLAST score
                   8.0e-49
E value
                   220
Match length
                   41
% identity
                   (U93215) maize transposon MuDR mudrA protein isolog
NCBI Description
                   [Arabidopsis thaliana] >gi_2880040 (AC002340) maize
                   transposon MuDR mudrA-like protein [Arabidopsis thaliana]
 Seq. No.
                   10127
                   17999 1.R1040
 Contig ID
                   LIB3106-043-Q1-K1-H5
 5'-most EST
                   10128
 Seq. No.
                   17999 2.R1040
 Contig ID
                   LIB3049-014-Q1-E1-E8
 5'-most EST
                   10129
 Seq. No.
                   18000 1.R1040
 Contig ID
                   epx701107566.hl
 5'-most EST
                   BLASTX
 Method
                   g2494100
 NCBI GI
                    699
 BLAST score
                    1.0e-73
 E value
```

Match length 259 49 % identity NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (NADP-GDH) NCBI Description (NAD(P)H-DEPENDENT GLUTAMATE DEHYDROGENASE) >gi 1772845 (U82240) NAD(P)H-dependent glutamate dehydrogenase

[Prevotella ruminicola]



```
Seq. No.
                  10130
                  18002 1.R1040
Contig ID
                  LIB3049-002-Q1-E1-F4
5'-most EST
Seq. No.
                  10131
                  18006 1.R1040
Contig ID
5'-most EST
                  uC-qmropic112c12b1
                  BLASTX
Method
NCBI GI
                  g1246403
BLAST score
                  291
                  8.0e-26
E value
Match length
                  151
                  44
% identity
NCBI Description (X94698) TINY [Arabidopsis thaliana] >gi_3406035 (AC005405)
                  TINY [Arabidopsis thaliana]
Seq. No.
                  10132
                  18012 1.R1040
Contig ID
                  LIB3170-022-Q1-K1-F8
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2842490
BLAST score
                  406
E value
                  8.0e-40
Match length
                  83
% identity
                  88
NCBI Description (AL021749) heat-shock protein [Arabidopsis thaliana]
Seq. No.
                  10133
                  18014 1.R1040
Contig ID
5'-most EST
                  kmv700737783.h1
                  BLASTX
Method
                  g3176669
NCBI GI
BLAST score
                  187
                   6.0e-14
E value
Match length
                  102
% identity
NCBI Description (AC004393) End is cut off. [Arabidopsis thaliana]
Seq. No.
                   10134
                   18016 1.R1040
Contig ID
                  LIB3107-013-Q1-K1-E5
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3461820
BLAST score
                   260
                  2.0e-22
E value
                  64
Match length
                  69
% identity
NCBI Description (AC004138) unknown protein [Arabidopsis thaliana]
                   10135
Seq. No.
Contig ID
                   18016 2.R1040
5'-most EST
                  LIB3109-020-Q1-K1-G3
Method
                  BLASTX
```

 NCBI GI
 g3461820

 BLAST score
 260

 E value
 3.0e-22

```
Match length
                  64
                   69
% identity
NCBI Description
                  (AC004138) unknown protein [Arabidopsis thaliana]
Seq. No.
                  10136
                  18022 1.R1040
Contig ID
5'-most EST
                  LIB3106-049-Q1-K1-B11
Seq. No.
                  10137
                  18028 1.R1040
Contig ID
```

LIB3049-013-Q1-E1-H3

 Seq. No.
 10138

 Contig ID
 18031_1.R1040

 5'-most EST
 uC-gmflminsoy071b04b2

 Method
 BLASTX

Method BLASTX
NCBI GI g2916917
BLAST score 633
E value 1.0e-80
Match length 373
% identity 44

5'-most EST

NCBI Description (AL022004) hypothetical protein Rv0858c [Mycobacterium

tuberculosis]

Seq. No. 10139

Contig ID 18036_1.R1040

5'-most EST LIB3049-014-Q1-E1-A11

Seq. No. 10140

Contig ID 18040_1.R1040

5'-most EST LIB $30\overline{4}$ 9-014-Q1-E1-A6

Seq. No. 10141

Contig ID 18041 1.R1040 5'-most EST crh700854610.h1

Method BLASTX
NCBI GI g3482974
BLAST score 1155
E value 1.0e-127
Match length 323
% identity 70

NCBI Description (AL031369) ATP-dependent Clp proteinase-like protein

[Arabidopsis thaliana]

Seq. No. 10142

Contig ID 18042 1.R1040

5'-most EST LIB3049-014-Q1-E1-A8

Method BLASTX
NCBI GI 94558659
BLAST score 456
E value 2.0e-45
Match length 118
% identity 69

NCBI Description (AC007063) unknown protein [Arabidopsis thaliana]

Seq. No. 10143

Contig ID 18047_1.R1040



```
5'-most EST
                  LIB3049-013-Q1-E1-F6
                  10144
Seq. No.
                  18050 1.R1040
Contig ID
5'-most EST
                  LIB3049-013-Q1-E1-F9
Method
                  BLASTX
NCBI GI
                  g2492772
BLAST score
                  527
E value
                  5.0e-54
Match length
                  104
% identity
                  93
                  GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE (FDH)
NCBI Description
                  (FALDH) (GSH-FDH) >gi 1498024 (U63931)
                  glutathione-dependent formaldehyde dehydrogenase
                  [Arabidopsis thaliana]
                  10145
Seq. No.
Contig ID
                  18052 1.R1040
5'-most EST
                  LIB3049-013-Q1-E1-G9
                  10146
Seq. No.
Contig ID
                  18062 1.R1040
5'-most EST
                  LIB3106-039-Q1-K1-D3
Method
                  BLASTX
NCBI GI
                  g1001603
BLAST score
                  187
                  6.0e-20
E value
                  117
Match length
% identity
                  50
NCBI Description
                 (D64000) hypothetical protein [Synechocystis sp.]
                  10147
Seq. No.
                  18066 1.R1040
Contig ID
5'-most EST
                  leu701146028.h1
Method
                  BLASTN
NCBI GI
                  g3869069
BLAST score
                  34
E value
                   6.0e-09
                  167
Match length
                  88
% identity
NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:
                  MEB5, complete sequence [Arabidopsis thaliana]
Seq. No.
                  10148
                  18066 2.R1040
Contig ID
5'-most EST
                  jC-gmle01810024b06d1
Seq. No.
                  10149
Contig ID
                  18066 3.R1040
```

5'-most EST LIB3170-065-Q1-K1-B5

Seq. No. 10150

Contig ID 18066_4.R1040 5'-most EST kmv700742292.h1

Seq. No. 10151

Contig ID 18068 1.R1040

Contig ID

5'-most EST



```
5'-most EST
                     uC-gmropic025h12b1
                     BLASTX
  Method
  NCBI GI
                     q984756
  BLAST score
                     761
                     1.0e-80
  E value
  Match length
                     396
                     44
  % identity
  NCBI Description (Z54153) chilling-inducible protein [Oryza sativa]
  Seq. No.
                     10152
                     18068 2.R1040
. Contig ID
  5'-most EST
                     qsv701054946.h1
                     BLASTX
  Method
                     g984756
  NCBI GI
  BLAST score
                     218
  E value
                     1.0e-17
  Match length
                     67
                     58
  % identity
  NCBI Description (Z54153) chilling-inducible protein [Oryza sativa]
  Seq. No.
                     10153
  Contig ID
                     18068 3.R1040
  5'-most EST
                     zhf700964316.h1
                     BLASTX
  Method
  NCBI GI
                     q984756
  BLAST score
                     333
  E value
                     5.0e-31
  Match length
                     140
                     49
  % identity
  NCBI Description (Z54153) chilling-inducible protein [Oryza sativa]
                     10154
  Seq. No.
                     18075 1.R1040
  Contig ID
  5'-most EST
                     jC-gmst02400025c10a1
                     BLASTX
  Method
  NCBI GI
                     q3420054
  BLAST score
                     219
  E value
                     1.0e-17
  Match length
                     125
   % identity
                     42
  NCBI Description (AC004680) unknown protein [Arabidopsis thaliana]
                     10155
   Seq. No.
                     18076 1.R1040
  Contig ID
   5'-most EST
                     LIB3051-084-Q1-K1-E8
                     10156
  Seq. No.
                     18076 2.R1040
   Contig ID
   5'-most EST
                     LIB3107-080-Q1-K1-D11
  Seq. No.
                     10157
   Contig ID
                     18081 1.R1040
   5'-most EST
                     LIB3049-013-Q1-E1-C2
                     10158
   Seq. No.
                     18084 1.R1040
```

LIB3052-014-Q1-N1-G9



Method BLASTX
NCBI GI g3123908
BLAST score 657
E value 1.0e-68
Match length 171
% identity 67

NCBI Description (AF038392) pre-mRNA splicing factor [Homo sapiens]

Seq. No. 10159

Contig ID 18086 1.R1040 5'-most EST gsv701044694.h1

Seq. No. 10160

Contig ID 18090_1.R1040

5'-most EST LIB3106-026-Q1-K1-H5

Method BLASTX
NCBI GI g2213594
BLAST score 581
E value 5.0e-60
Match length 161
% identity 65

NCBI Description (AC000348) T7N9.14 [Arabidopsis thaliana]

Seq. No. 10161

Contig ID 18092 1.R1040

5'-most EST LIB3049-013-Q1-E1-D2

Seq. No. 10162

Contig ID 18098 1.R1040

5'-most EST LIB3049-013-Q1-E1-C12

Seq. No. 10163

Contig ID 18100_1.R1040 5'-most EST sat701003273.h1

Method BLASTX
NCBI GI g3785998
BLAST score 426
E value 8.0e-42
Match length 121
% identity 61

NCBI Description (AC005499) unknown protein [Arabidopsis thaliana]

Seq. No. 10164

Contig ID 18100_2.R1040

5'-most EST LIB3170-024-Q1-K1-F2

Method BLASTX
NCBI GI g3785998
BLAST score 335
E value 3.0e-31
Match length 94
% identity 64

NCBI Description (AC005499) unknown protein [Arabidopsis thaliana]

Seq. No. 10165

Contig ID 18103 1.R1040 5'-most EST trc700562330.h1

Method BLASTX



q3273202 NCBI GI BLAST score 414 E value 2.0e-40 Match length 106 76 % identity NCBI Description (AB010918) responce reactor4 [Arabidopsis thaliana] Seq. No. 10166 18103 2.R1040 Contig ID 5'-most EST LIB3051-068-Q1-K1-E2 Method BLASTX q3273202

Method BLASTX
NCBI GI g3273202
BLAST score 179
E value 1.0e-12
Match length 162
% identity 31

NCBI Description (AB010918) responce reactor4 [Arabidopsis thaliana]

Seq. No. 10167

Contig ID 18106 1.R1040

5'-most EST LIB3170-023-Q1-K1-F8

Method BLASTN
NCBI GI g3510347
BLAST score 45
E value 4.0e-16
Match length 182
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:

MSJ11, complete sequence [Arabidopsis thaliana]

Seq. No. 10168

Contig ID 18109_1.R1040 5'-most EST uaw700663343.h1

Seq. No. 10169

Contig ID 18112 1.R1040

5'-most EST LIB3049-013-Q1-E1-C1

Seq. No. 10170

Contig ID 18114 1.R1040

5'-most EST jC-gmle01810005b01a1

Method BLASTX
NCBI GI 94006826
BLAST score 448
E value 2.0e-44
Match length 148
% identity 59

NCBI Description (AC005970) unknown protein [Arabidopsis thaliana]

Seq. No. 10171

Contig ID 18116 1.R1040 5'-most EST leu701151307.h1

Method BLASTX
NCBI GI g2645699
BLAST score 182
E value 4.0e-13
Match length 92

% identity

(AF031933) glycine-rich RNA-binding protein [Euphorbia NCBI Description

esula]

Seq. No.

10172

Contig ID 5'-most EST

18117 1.R1040 leu701151765.h1

Seq. No.

10173

Contig ID 5'-most EST 18122 1.R1040 epx701105809.h1

Seq. No.

10174

Contig ID

18129 1.R1040

5'-most EST

LIB3049-013-Q1-E1-A12

Seq. No.

10175

Contig ID 5'-most EST

18129 2.R1040 pmv700891366.hl

Seq. No.

10176

Contig ID 5'-most EST

18130 1.R1040 kl1701205527.h1

Seq. No.

10177

Contig ID

18136 1.R1040

5'-most EST

LIB3170-022-Q1-K1-E10 BLASTX

Method NCBI GI

g4415912

BLAST score

169

E value

6.0e-12

Match length

65

% identity

62

NCBI Description

(AC006282) putative protease [Arabidopsis thaliana]

Seq. No.

10178

Contia ID 5'-most EST 18138 1.R1040 pxt700944466.hl

10179

Seq. No. Contig ID

18139 1.R1040

5'-most EST

LIB3170-070-Q1-K2-B6

Method

BLASTX

NCBI GI BLAST score g3763916 293

E value

3.0e-26

Match length

103

% identity

51

NCBI Description

(AC004450) unknown protein [Arabidopsis thaliana]

>gi 4531439 gb AAD22124.1 AC006224 6 (AC006224) unknown

protein [Arabidopsis thaliana]

Seq. No.

10180

Contig ID

18139 2.R1040

5'-most EST

hrw701061944.hl

Method

BLASTX

NCBI GI

g3763916



BLAST score 158
E value 1.0e-10
Match length 79
% identity 39

NCBI Description (AC004450) unknown protein [Arabidopsis thaliana]

>gi_4531439_gb_AAD22124.1_AC006224_6 (AC006224) unknown

protein [Arabidopsis thaliana]

Seq. No. 10181

Contig ID 18142_1.R1040 5'-most EST pmv700890296.h1

Seq. No. 10182

Contig ID 18146 1.R1040

5'-most EST LIB3049-012-Q1-E1-G7

Seq. No. 10183

Contig ID 18148_1.R1040

5'-most EST LIB3170-046-Q1-J1-D1

Method BLASTX
NCBI GI g3660469
BLAST score 1827
E value 0.0e+00
Match length 410
% identity 87

NCBI Description (AJ001808) succinyl-CoA-ligase beta subunit [Arabidopsis

thaliana] >gi 4512693 gb AAD21746.1 (AC006569)

succinyl-CoA ligase beta subunit [Arabidopsis thaliana]

Seq. No. 10184

Contig ID 18148_2.R1040 5'-most EST sat701012817.h1

Method BLASTX
NCBI GI g3660469
BLAST score 423
E value 1.0e-41
Match length 97
% identity 82

NCBI Description (AJ001808) succinyl-CoA-ligase beta subunit [Arabidopsis

thaliana] >gi 4512693 gb AAD21746.1 (AC006569)

succinyl-CoA ligase beta subunit [Arabidopsis thaliana]

Seq. No. 10185

Contig ID 18149_1.R1040

5'-most EST LIB3049-012-Q1-E1-D5

Seq. No. 10186

Contig ID 18153 1.R1040 5'-most EST gsv701046028.h1

Seq. No. 10187

Contig ID 18156_1.R1040 5'-most EST ncj700975168.h1

Method BLASTX
NCBI GI g3096931
BLAST score 362
E value 3.0e-34

```
Match length
                  111
% identity
                  59
                  (AL023094) putative ribosomal protein S16 [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  10188
Contig ID
                  18157 1.R1040
                  jex700907657.hl
5'-most EST
                  BLASTX
Method
                  q4455199
NCBI GI
BLAST score
                  1108
E value
                  1.0e-121
Match length
                  283
% identity
                  78
                 (AL035440) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  10189
Contig ID
                  18157 2.R1040
5'-most EST
                  eep700868548.hl
Method
                  BLASTX
NCBI GI
                  q4522009
BLAST score
                  389
                  9.0e-38
E value
Match length
                  127
% identity
                  54
NCBI Description
                  (AC007069) unknown protein [Arabidopsis thaliana]
Seq. No.
                  10190
                  18158 1.R1040
Contig ID
5'-most EST
                  zzp700835245.h1
Method
                  BLASTX
NCBI GI
                  g3204101
```

251 BLAST score 3.0e-21 E value

102 Match length % identity 58

NCBI Description (AJ006760) hypothetical protein [Cicer arietinum]

Seq. No. 10191 18158 2.R1040 Contig ID 5'-most EST gsv701045938.h1

Seq. No. 10192

18166 1.R1040 Contig ID LIB3109-016-Q1-K3-B5

5'-most EST Method BLASTX

NCBI GI q4567215 274 BLAST score 3.0e-24 E value 113 Match length % identity

NCBI Description (AC007113) putative presenilin [Arabidopsis thaliana]

Seq. No. 10193

Contig ID 18169 1.R1040 5'-most EST q5753705 Method BLASTX



g2245101 NCBI GI BLAST score 517 2.0e-52 E value 144 Match length 71 % identity NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana] Seq. No. 10194 18169 3.R1040 Contig ID LIB3170-022-Q1-K1-E5 5'-most EST

BLASTX Method g2245101 NCBI GI BLAST score 332 E value 6.0e-31 Match length 72

86 % identity

NCBI Description (297343) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10195

Contig ID 18169 4.R1040 5'-most EST gsv701047470.h1

Seq. No. 10196

18169 5.R1040 Contig ID 5'-most EST eep700865796.hl

Method BLASTX q2245101 NCBI GI BLAST score 231 3.0e-19 E value Match length 49

90 % identity

NCBI Description (297343) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10197

Contig ID 18170 1.R1040 5'-most EST jex700906004.h1

Method BLASTX g1808656 NCBI GI BLAST score 1381 1.0e-153 E value 311 Match length 83 % identity

NCBI Description (Y10804) Ubiquitin activating enzyme El [Nicotiana tabacum]

Seq. No. 10198

18171 1.R1040 Contig ID

5'-most EST LIB3049-012-Q1-E1-C10

10199 Seq. No.

Contig ID 18179 1.R1040

5'-most EST LIB3049-012-Q1-E1-B5

10200 Seq. No.

18180 1.R1040 Contig ID

5'-most EST LIB3139-094-P1-N1-H7

Method BLASTX NCBI GI g128405

1789



BLAST score 457 E value 3.0e-45Match length 166 % identity 57

NODULIN 21 (N-21) >gi_99942_pir__S08632 nodulin-21 - soybean >gi_18694_emb_CAA34506_ (X16488) nodulin-21 (AA NCBI Description

1-201) [Glycine max]

Seq. No.

10201

Contig ID

18183 1.R1040

5'-most EST

LIB3049-012-Q1-E1-A10

Method BLASTX NCBI GI g3096939 BLAST score 333 E value 6.0e-31 Match length 87 % identity 72

NCBI Description (AL023094) putative protein [Arabidopsis thaliana]

Seq. No.

10202

Contig ID 5'-most EST 18183 2.R1040 leu701150481.h1

Method BLASTN NCBI GI g3641835 BLAST score 39 E value 2.0e-12 Match length 153

85 % identity

NCBI Description Arabidopsis thaliana DNA chromosome 4, BAC clone T4L20

(ESSAII project)

Seq. No.

10203

Contig ID

18188 1.R1040

5'-most EST

LIB3049-012-Q1-E1-A5

Seq. No.

10204

Contig ID 5'-most EST 18188 2.R1040 hrw701060104.h1

Seq. No.

10205

Contig ID

18192 1.R1040

5'-most EST

LIB3109-038-Q1-K1-A12

Seq. No.

10206

Contig ID

18194 1.R1040

5'-most EST

LIB3056-002-Q1-B1-E5

Method BLASTX NCBI GI g3327957 BLAST score 209 E value 3.0e-16 Match length 91 46 % identity

NCBI Description

(AF060490) TLS-associated protein TASR-2 [Mus musculus] >gi 3327976 (AF067730) TLS-associated protein TASR-2 [Homo

sapiens]

Seq. No.

10207



Contig ID 18207 1.R1040

5'-most EST uC-gmrominsoy090a03b1

Method BLASTX
NCBI GI g2281637
BLAST score 304
E value 4.0e-27
Match length 65
% identity 88

NCBI Description (AF003099) AP2 domain containing protein RAP2.6

[Arabidopsis thaliana]

Seq. No. 10208

Contig ID 18207_2.R1040

5'-most EST LIB3049-011-Q1-E1-H10

Seq. No. 10209

Contig ID 18207_3.R1040 5'-most EST uC-gmropic0001c10b1

Seq. No. 10210

Contig ID 18207_4.R1040 5'-most EST jsh701067143.h1

Seq. No. 10211

Contig ID 18207_6.R1040

5'-most EST LIB3051-050-Q1-K1-F3

Seq. No. 10212

Contig ID 18207 9.R1040

5'-most EST LIB3139-101-P1-N1-G9

Seq. No. 10213

Contig ID 18216 1.R1040 5'-most EST hrw701062463.h1

Seq. No. 10214

Contig ID 18216_2.R1040 5'-most EST uaw700664293.h1

Seq. No. 10215

Contig ID 18217_1.R1040

5'-most EST LIB3049-011-Q1-E1-E4

Method BLASTX
NCBI GI g2827715
BLAST score 539
E value 2.0e-55
Match length 128
% identity 12

NCBI Description (AL021684) receptor protein kinase - like protein

[Arabidopsis thaliana]

Seq. No. 10216

Contig ID 18219 1.R1040

5'-most EST LIB3170-023-Q1-K1-E4

Seq. No. 10217

Contig ID 18220 1.R1040



```
LIB3170-023-Q1-K1-E5
5'-most EST
                  10218
Seq. No.
                  18221 1.R1040
Contig ID
5'-most EST
                  LIB3049-011-Q1-E1-F11
                  10219
Seq. No.
                  18223 1.R1040
Contig ID
                  LIB3049-011-Q1-E1-F3
5'-most EST
                  10220
Seq. No.
                  18231 1.R1040
Contig ID
                  jC-qmst02400060b11a1
5'-most EST
                  BLASTX
Method
                  g4510406
NCBI GI
BLAST score
                  550
E value
                  3.0e-56
Match length
                  150
% identity
                  (AC006587) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                  10221
Seq. No.
                  18231 2.R1040
Contig ID
                  pmv700892967.h1
5'-most EST
                  BLASTX
Method
                   q3367520
NCBI GI
BLAST score
                   440
                   2.0e-47
E value
                   174
Match length
% identity
                   60
                   (AC004392) Similar to protein kinase APK1A,
NCBI Description
                   tyrosine-serine-threonine kinase gb_D12522 from A.
                   thaliana. [Arabidopsis thaliana]
                   10222
Seq. No.
                   18231 3.R1040
Contig ID
                   epx701104778.hl
5'-most EST
                   10223
Seq. No.
                   18233 1.R1040
Contig ID
                   LIB3049-011-Q1-E1-E3
5'-most EST
                   BLASTX
Method
                   g2462753
NCBI GI
                   876
BLAST score
                   1.0e-94
E value
                   204
Match length
                   77
% identity
                   (AC002292) putative polygalacturonase [Arabidopsis
NCBI Description
                   thaliana]
                   10224
Seq. No.
                   18236 1.R1040
Contig ID
                   LIB3049-011-Q1-E1-C6
5'-most EST
                   10225
Seq. No.
```

1792

18238 1.R1040

uC-gmflminsoy055f12b1

Contig ID

5'-most EST



Method BLASTX NCBI GI q2288887 BLAST score 642 E value 5.0e-67 Match length 167 73 % identity

(Y14325) mevalonate diphosphate decarboxylase [Arabidopsis NCBI Description thaliana] >gi_3250736_emb_CAA76803_ (Y17593) mevalonate

diphosphate decarboxylase [Arabidopsis thaliana]

>gi 3786002 (AC005499) mevalonate diphosphate decarboxylase

[Arabidopsis thaliana]

Seq. No. 10226

18241 1.R1040 Contig ID

5'-most EST LIB3049-011-Q1-E1-D11

Method BLASTX NCBI GI q1737492 BLAST score 1108 E value 1.0e-121 407 Match length % identity

NCBI Description (U81318) poly(A)-binding protein [Triticum aestivum]

Seq. No. 10227

Contig ID 18245 1.R1040

5'-most EST LIB3049-011-Q1-E1-D6

Method BLASTX NCBI GI g3913732 BLAST score 271 1.0e-23 E value Match length 86 % identity 62

NCBI Description

HYDROXYACYLGLUTATHIONE HYDROLASE MITOCHONDRIAL ISOZYME PRECURSOR (GLYOXALASE II) (GLX II) >gi_2570340 (U90928)

glyoxalase II mitochondrial isozyme [Arabidopsis thaliana]

Seq. No. 10228

Contig ID 18245 2.R1040 5'-most EST kl1701203505.h2

Seq. No. 10229

18246 1.R1040 Contig ID 5'-most EST dpv701099471.h1

10230 Seq. No.

18246 3.R1040 Contig ID 5'-most EST zhf700962535.h1

Seq. No. 10231

18254 1.R1040 Contig ID

5'-most EST LIB3049-011-Q1-E1-A8

Method BLASTX NCBI GI q3600039 BLAST score 232 2.0e-19 E value Match length 54 % identity 76



NCBI Description (AF080119) similar to Schizosaccharomyces pombe isp4 protein (GB:D14061) [Arabidopsis thaliana]

Seq. No. 10232

Contig ID 18256_1.R1040 5'-most EST xpa700792801.h1

Method BLASTX
NCBI GI g2809233
BLAST score 239
E value 3.0e-20
Match length 75
% identity 63

NCBI Description (AC002560) F21B7.2 [Arabidopsis thaliana]

Seq. No.

Contig ID 18260_1.R1040

5'-most EST LIB3170-081-Q1-K1-D3

10233

Method BLASTN
NCBI GI g170007
BLAST score 112
E value 9.0e-56
Match length 196
% identity 90

NCBI Description Soybean 18 kD late embryogenesis abundant (Lea) protein

mRNA, complete cds

Seq. No. 10234

Contig ID 18262_1.R1040 5'-most EST bth700845315.h1 Method BLASTX

Method BLASTX
NCBI GI g4512685
BLAST score 279
E value 1.0e-24
Match length 150
% identity 42

NCBI Description (AC006931) hypothetical protein [Arabidopsis thaliana]

>gi_4559325_gb_AAD22987.1_AC007087_6 (AC007087)

hypothetical protein [Arabidopsis thaliana]

Seq. No. 10235

Contig ID 18264_1.R1040 5'-most EST hrw701059985.h1

Method BLASTX
NCBI GI g3880215
BLAST score 156
E value 8.0e-10
Match length 142
% identity 32

NCBI Description (Z82053) cDNA EST yk302b12.3 comes from this gene

[Caenorhabditis elegans]

Seq. No. 10236

Contig ID 18265_1.R1040 5'-most EST uxk700673114.h1

Method BLASTX NCBI GI g4510345 BLAST score 463



6.0e-46 E value Match length 177 % identity 51 (AC006921) unknown protein [Arabidopsis thaliana] NCBI Description 10237 Seq. No. 18265 2.R1040 Contig ID 5'-most EST LIB3049-011-Q1-E1-C2 Method BLASTX NCBI GI g4510345 BLAST score 465 4.0e-46 E value 176 Match length 51 % identity NCBI Description (AC006921) unknown protein [Arabidopsis thaliana]

Seq. No. 10238

Contig ID 18265 3.R1040 rca701002474.h1 5'-most EST

Method BLASTX NCBI GI q4510345 BLAST score 456 5.0e-45 E value Match length 172 % identity

NCBI Description (AC006921) unknown protein [Arabidopsis thaliana]

Seq. No. 10239

Contig ID 18265 4.R1040 5'-most EST epx701106091.hl

Method BLASTX NCBI GI g4510345 BLAST score 189 E value 2.0e-15 55 Match length % identity 72

NCBI Description (AC006921) unknown protein [Arabidopsis thaliana]

Seq. No. 10240

18265 5.R1040 Contig ID 5'-most EST uC-qmropic099f11b1

Method BLASTX NCBI GI g4510345 229 BLAST score 5.0e-19 E value Match length 54 74 % identity

NCBI Description (AC006921) unknown protein [Arabidopsis thaliana]

Seq. No. 10241

18266 1.R1040 Contig ID 5'-most EST uC-gmropic018a05b1

BLASTX Method NCBI GI q2642154 BLAST score 526 E value 4.0e-55 Match length 216



% identity 58

NCBI Description (AC003000) unknown protein [Arabidopsis thaliana] >qi 3790595 (AF079186) RING-H2 finger protein RHC2a

[Arabidopsis thaliana]

Seq. No. 10242

Contig ID 18267_2.R1040 5'-most EST zzp700832279.h1

Seq. No. 10243

Contig ID 18267 3.R1040 5'-most EST zpv700757039.h1

Seq. No. 10244

Contig ID 18271_1.R1040 5'-most EST gsv701051833.h1

Method BLASTX
NCBI GI g2952433
BLAST score 625
E value 2.0e-65
Match length 146
% identity 75

NCBI Description (AF051135) putative ubiquitin activating enzyme E1

[Arabidopsis thaliana]

Seq. No. 10245

Contig ID 18272 1.R1040

5'-most EST LIB3049-010-Q1-E1-H4

Method BLASTX
NCBI GI g2980782
BLAST score 228
E value 1.0e-18
Match length 141
% identity 44

NCBI Description (AL022198) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10246

Contig ID 18290 1.R1040 5'-most EST wvk700680196.h2

Method BLASTX
NCBI GI g3979937
BLAST score 154
E value 4.0e-10
Match length 67
% identity 43

NCBI Description (AL034393) predicted using Genefinder; cDNA EST EMBL:D65716

comes from this gene; cDNA EST yk263b1.3 comes from this gene; cDNA EST yk263b1.5 comes from this gene; cDNA EST EMBL:D69251 comes from this gene [Caenorhabditis elegans]

Seq. No. 10247

Contig ID 18291_1.R1040

5'-most EST LIB3170-021-Q1-K1-D12

Seq. No. 10248

Contig ID 18293 1.R1040

5'-most EST LIB3170-022-Q1-K1-D12



Seq. No. 10249

Contig ID 18294_1.R1040
5'-most EST g5342465
Method BLASTX
NCBI GI g4416307
BLAST score 193
E value 3.0e-14

E value 3.06
Match length 203
% identity 27

NCBI Description (AF105716) hypothetical protein [Zea mays]

Seq. No. 10250

Contig ID 18301 1.R1040

5'-most EST LIB3049-010-Q1-E1-E1

Seq. No. 10251

Contig ID 18303 1.R1040

5'-most EST LIB3139-040-P1-N1-F3

Seq. No. 10252

Contig ID 18306 1.R1040

5'-most EST LIB3049-010-Q1-E1-E4

Seq. No. 10253

Contig ID 18311 1.R1040

5'-most EST LIB3051-037-Q1-K1-B4

Method BLASTX
NCBI GI g3402711
BLAST score 165
E value 3.0e-11
Match length 46

% identity 76

NCBI Description (AC004261) putative RNA-binding protein [Arabidopsis

thaliana]

Seq. No. 10254

Contig ID 18312_1.R1040

5'-most EST LIB3049-010-Q1-E1-D6

Seq. No. 10255

Contig ID 18314 1.R1040

5'-most EST LIB3138-015-Q1-N2-H4

Method BLASTX
NCBI GI g2342724
BLAST score 1185
E value 1.0e-130
Match length 331
% identity 71

NCBI Description (AC002341) unknown protein [Arabidopsis thaliana]

Seq. No. 10256

Contig ID 18314 2.R1040

5'-most EST jC-gmst02400069b04a1

Seq. No. 10257

Contig ID 18319 1.R1040



```
5'-most EST
                   uC-qmropic090e05b1
Seq. No.
                   10258
                   18319 2.R1040
Contig ID
5'-most EST
                   LIB3049-010-Q1-E1-C7
                   10259
Seq. No.
                   18325_1.R1040
Contig ID
5'-most EST
                   LIB3093-023-Q1-K1-D10
Method
                   BLASTX
NCBI GI
                   q3212866
BLAST score
                   430
E value
                   8.0e-42
                   267
Match length
                   37
% identity
NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]
                   10260
Seq. No.
                   18325 2.R1040
Contig ID
5'-most EST
                   k1170\overline{1}214357.h1
                   BLASTX
Method
NCBI GI
                   a3212866
BLAST score
                   231
E value
                   4.0e-28
Match length
                   205
% identity
                   38
NCBI Description (AC004005) unknown protein [Arabidopsis thaliana]
Seq. No.
                   10261
                   18333 1.R1040
Contig ID
5'-most EST
                   leu701146986.h1
                   BLASTX
Method
NCBI GI
                   g3738331
BLAST score
                   322
                   9.0e-30
E value
Match length
                   72
% identity
                   74
NCBI Description (AC005170) unknown protein [Arabidopsis thaliana]
                   10262
Seq. No.
                   18340 1.R1040
Contig ID
5'-most EST
                   k11701208052.h1
                   BLASTX
Method
NCBI GI
                   g1176658
                   227
BLAST score
                   3.0e-18
E value
```

80 Match length 53 % identity

NCBI Description HYPOTHETICAL 200.6 KD PROTEIN B0228.2 IN CHROMOSOME II

>gi 726363 (U23168) No definition line found

[Caenorhabditis elegans]

10263 Seq. No.

Contig ID 18340 2.R1040

5'-most EST LIB3049-010-Q1-E1-B4

Seq. No. 10264



Contig ID 18340_3.R1040 5'-most EST zsg701123960.h1

Seq. No. 10265

Contig ID 18346_1.R1040 5'-most EST pmv700890645.h1

Seq. No. 10266

Contig ID 18349 1.R1040

5'-most EST LIB3170-022-Q1-K1-D4

Seq. No. 10267

Contig ID 18350 1.R1040

5'-most EST LIB30 $\overline{4}$ 9-009-Q1-E1-G7

Method BLASTX
NCBI GI g1800223
BLAST score 310
E value 2.0e-28
Match length 126
% identity 55

NCBI Description (U65018) mannosyltransferase [Dictyostelium discoideum]

Seq. No.

10268

Contig ID 18353 1.R1040

5'-most EST LIB3049-009-Q1-E1-H11

Seq. No. 10269

Contig ID 18355_1.R1040 5'-most EST uxk700668561.h1

Method BLASTX
NCBI GI g1742187
BLAST score 182
E value 4.0e-13
Match length 205
% identity 27

NCBI Description (D90771) ORF ID: 0260#14; similar to [SwissProt Accession

Number P11666] [Escherichia coli] >gi_1742198_dbj_BAA14933_(D90772) ORF_ID:o260#14; similar to [SwissProt Accession Number P11666] [Escherichia coli] >gi_1787591 (AE000231)

orf, hypothetical protein [Escherichia coli]

Seq. No. 10270

Contig ID 18356_1.R1040

5'-most EST LIB3049-009-Q1-E1-H3

Seq. No. 10271

Contig ID 18360_1.R1040 5'-most EST fC-gmf1700898745a1

Seq. No. 10272

Contig ID 18363_1.R1040 5'-most EST ssr700557079.h1

Method BLASTX
NCBI GI g4415912
BLAST score 180
E value 4.0e-13
Match length 70



% identity 50
NCBI Description (AC006282) putative protease [Arabidopsis thaliana]

Seq. No. 10273

Contig ID 18364_1.R1040

5'-most EST LIB3092-002-Q1-K1-H9

Method BLASTX
NCBI GI g3201680
BLAST score 640
E value 6.0e-67
Match length 163
% identity 72

NCBI Description (AF060941) extra-large G-protein [Arabidopsis thaliana]

Seq. No. 10274

Contig ID 18364 2.R1040

5'-most EST jC-qmro02910061q03a1

Method BLASTX
NCBI GI g3242709
BLAST score 602
E value 2.0e-62
Match length 197
% identity 63

NCBI Description (AC003040) putative guanine nucleotide-binding protein

[Arabidopsis thaliana]

Seq. No. 10275

Contig ID 18369_1.R1040 5'-most EST awf700838918.h1

Method BLASTX
NCBI GI g2244964
BLAST score 376
E value 6.0e-36
Match length 204
% identity 45

NCBI Description (Z97340) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10276

Contig ID 18370_1.R1040 5'-most EST gsv701046362.h1

Method BLASTX
NCBI GI g2982247
BLAST score 258
E value 7.0e-22
Match length 109
% identity 43

NCBI Description (AF051206) probable thioredoxin H [Picea mariana]

Seq. No. 10277

Contig ID 18377 1.R1040

5'-most EST LIB3049-009-Q1-E1-G1

Method BLASTX
NCBI GI g4115377
BLAST score 343
E value 6.0e-33
Match length 118
% identity 62

1800



```
(AC005967) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  10278
Contig ID
                  18378 1.R1040
5'-most EST
                  LIB3049-009-Q1-E1-G10
Seq. No.
                  10279
                  18380 1.R1040
Contig ID
5'-most EST
                  LIB3049-009-Q1-E1-G12
Seq. No.
                  10280
Contig ID
                  18387 1.R1040
5'-most EST
                  LIB3049-009-Q1-E1-D4
Seq. No.
                  10281
Contig ID
                  18389 1.R1040
5'-most EST
                  vzy700754587.hl
```

Method BLASTX NCBI GI q3559811 BLAST score 1491 E value 1.0e-166 Match length 401

% identity

(AJ010735) gr1-protein [Arabidopsis thaliana] NCBI Description

10282 Seq. No.

Contig ID 18390 1.R1040

5'-most EST LIB3065-006-Q1-N1-H12

Method BLASTX NCBI GI q416922 BLAST score 348 E value 7.0e-33 Match length 96 76 % identity

DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE (DUTPASE) NCBI Description

(DUTP PYROPHOSPHATASE) (P18) >gi 282947_pir__JQ1599 dUTP

pyrophosphatase (EC 3.6.1.23) - Tomato

>gi 251897 bbs 109276 (S40549) deoxyuridine triphosphatase,

dUTPase, P18 (EC 3.6.1.23) [tomatoes, Tint Tim cultivar

LA154, Peptide, 169 aa] [Lycopersicon esculentum]

10283 Seq. No.

Contig ID 18391 1.R1040

5'-most EST LIB3049-009-Q1-E1-D8

Method BLASTX NCBI GI g3142297 BLAST score 363 1.0e-34 E value Match length 96 74 % identity

(AC002411) Contains similarity to serine/threonine protein NCBI Description

phosphatase gb X83099 from S. cerevisiae. [Arabidopsis

thaliana]

10284 Seq. No.

Contig ID 18395 1.R1040

5'-most EST LIB3170-022-Q1-K1-D1



```
10285
Seq. No.
                  18396 1.R1040
Contig ID
                  LIB3049-009-Q1-E1-E2
5'-most EST
Method
                  BLASTN
NCBI GI
                  q1806143
                  142
BLAST score
                   6.0e-74
E value
                   274
Match length
% identity
                   88
NCBI Description M.sativa mRNA for cdc2 kinase homologue, cdc2MsE
                   10286
Seq. No.
                   18414 1.R1040
Contig ID
                   zhf700951883.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g1944575
BLAST score
                   537
                   4.0e-55
E value
                   133
Match length
                   76
% identity
                  (Z94058) pectinesterase [Lycopersicon esculentum]
NCBI Description
                   10287
Seq. No.
                   18415 1.R1040
Contig ID
                   LIB3170-023-Q1-K1-D4
5'-most EST
                   10288
Seq. No.
                   18420 1.R1040
Contig ID
5'-most EST
                   bth700845450.h1
                   10289
Seq. No.
Contig ID
                   18422 1.R1040
                   LIB3170-023-Q1-K1-D10
5'-most EST
                   10290
Seq. No.
Contig ID
                   18425 1.R1040
                   pmv70\overline{0}889715.h1
5'-most EST
Method
                   BLASTX
                   g1352881
NCBI GI
BLAST score
                   154
                   5.0e-10
E value
Match length
                   95
% identity
                   36
                   HYPOTHETICAL 15.7 KD PROTEIN IN NUP85-SSC1 INTERGENIC
NCBI Description
                   REGION >gi 1078266 pir__S57063 probable membrane protein
                   YJR044c - yeast (Saccharomyces cerevisiae)
                   >gi 1015699_emb_CAA89572_ (Z49544) ORF YJR044c
                   [Saccharomyces cerevisiae] >gi_1197072 (L36344) ORF;
                   putative [Saccharomyces cerevisiae]
```

Contig ID 18425_2.R1040 5'-most EST ncj700980410.h1

Seq. No. 10292

Contig ID 18426_1.R1040

NCBI GI



```
uC-qmflminsoy005c06b1
5'-most EST
                  10293
Seq. No.
                  18426 2.R1040
Contig ID
                  uC-gmrominsoy099e12b1
5'-most EST
                  10294
Seq. No.
                  18427 1.R1040
Contig ID
                  fua701040106.hl
5'-most EST
                  BLASTX
Method
                  q3763940
NCBI GI
                  535
BLAST score
                  1.0e-54
E value
                  157
Match length
% identity
                  71
                  (AC004450) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  10295
Seq. No.
                  18429 1.R1040
Contig ID
                  LIB3170-024-Q1-K1-D7
5'-most EST
                  10296
Sea. No.
                  18436 1.R1040
Contig ID
                  LIB3109-006-Q1-K1-H9
5'-most EST
Seq. No.
                   10297
                   18438 1.R1040
Contig ID
                  LIB3049-008-Q1-E1-G2
5'-most EST
                   10298
Seq. No.
                   18439 1.R1040
Contig ID
                   LIB3049-008-Q1-E1-G3
5'-most EST
                   BLASTX
Method
                   g2760835
NCBI GI
                   298
BLAST score
                   7.0e-27
E value
                   134
Match length
                   46
% identity
                  (AC003105) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10299
Seq. No.
Contig ID
                   18441 1.R1040
                   wrg700790675.h2
5'-most EST
Method
                   BLASTX
                   g1764100
NCBI GI
                   290
BLAST score
                   1.0e-30
E value
                   94
Match length
                   72
% identity
                   (U81805) GDP-D-mannose-4,6-dehydratase [Arabidopsis
NCBI Description
                   thaliana]
                   10300
Seq. No.
                   18442 1.R1040
Contig ID
                   LIB3049-008-Q1-E1-G6
5'-most EST
Method
                   BLASTX
```

1803

g3152605



BLAST score 2.0e-26 E value 154 Match length % identity NCBI Description (AC004482) hypothetical protein [Arabidopsis thaliana] 10301 Seq. No. 18444 1.R1040 Contig ID LIB3049-008-Q1-E1-G8 5'-most EST BLASTX Method NCBI GI q3355465 BLAST score 305 7.0e-28 E value 125 Match length % identity (AC004218) putative Ser/Thr protein kinase [Arabidopsis NCBI Description thaliana] 10302 Seq. No. 18445 1.R1040 Contig ID asn701138856.h1 5'-most EST BLASTX Method q1653775 NCBI GI 220 BLAST score 1.0e-17 E value 117 Match length % identity 46 (D90916) thiol:disulfide interchange protein DsbD NCBI Description [Synechocystis sp.] Seq. No. 10303 Contig ID 18446 1.R1040 LIB3138-080-P1-N1-C1 5'-most EST Method BLASTX NCBI GI g2245378 BLAST score 658 E value 1.0e-77 198 Match length % identity 71 (U83245) auxin response factor 1 [Arabidopsis thaliana] NCBI Description 10304 Seq. No. 18446 2.R1040 Contig ID LIB3049-006-Q1-E1-B8 5'-most EST 10305 Seq. No. 18447 1.R1040 Contig ID LIB3170-036-Q1-K2-B8 5'-most EST BLASTX Method g3164222 NCBI GI BLAST score 252 E value 1.0e-21

Match length 64 67 % identity

(AB008518) RMA1 [Arabidopsis thaliana] >gi 4206205 NCBI Description (AF071527) RMA1 RING zinc finger protein [Arabidopsis

thaliana]

1804



```
10306
Seq. No.
                                                                                 .....
                   18447 2.R1040
Contig ID
                   zhf70\overline{0}962429.h1
5'-most EST
Seq. No.
                   10307
                   18449 1.R1040
Contig ID
5'-most EST
                   LIB3170-024-Q1-K1-D6
                   BLASTX
Method
                   q125606
NCBI GI
BLAST score
                   197
E value
                   3.0e-15
Match length
                   63
                   63
% identity
                   PYRUVATE KINASE, CYTOSOLIC ISOZYME >gi_100463 pir__S12248
NCBI Description
                   pyruvate kinase (EC 2.7.1.40) - potato
                   >gi 22576 emb CAA37727_ (X53688) pyruvate kinase [Solanum
                   tuberosum]
                   10308
Seq. No.
                   18450 1.R1040
Contig ID
5'-most EST
                   gsv701047939.hl
                   BLASTX
Method
                   q2760830
NCBI GI
                   1576
BLAST score
                   1.0e-176
E value
Match length
                   382
% identity
                   (AC003105) putative beta-ketoacyl-CoA synthase [Arabidopsis
NCBI Description
                   thaliana]
                   10309
Seq. No.
                   18460 1.R1040
Contig ID
                   LIB3170-023-Q1-K1-D8
5'-most EST
Seq. No.
                   10310
Contig ID
                   18461 1.R1040
                   smc70\overline{0}749767.h1
5'-most EST
Seq. No.
                   10311
Contig ID
                   18463 1.R1040
5'-most EST
                   LIB3109-022-Q1-K1-D8
Method
                   BLASTX
                   g3420054
NCBI GI
                   269
BLAST score
                   2.0e-23
E value
                   183
Match length
% identity
                   39
                   (AC004680) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10312
Seq. No.
                   18464 1.R1040
```

Contig ID

5'-most EST

LIB3049-008-Q1-E1-F3

Seq. No.

10313

Contig ID

18469 1.R1040

5'-most EST

pmv700891633.h1

1805

% identity



```
BLASTX
Method
                  g3850063
NCBI GI
                  787
BLAST score
                  1.0e-83
E value
                  482
Match length
                  35
% identity
NCBI Description (AJ223830) ARE1 [Rattus norvegicus]
                  10314
Seq. No.
                  18478 1.R1040
Contig ID
                  LIB3170-024-Q1-J1-D1
5'-most EST
                  BLASTX
Method
                  q3834322
NCBI GI
                  180
BLAST score
E value
                   4.0e-13
Match length
                   46
                   80
% identity
NCBI Description (AC005679) EST gb_R30300 comes from this gene. [Arabidopsis
                  thaliana]
                   10315
Seq. No.
                  18481 1.R1040
Contig ID
                  LIB3170-022-Q1-K1-C11
5'-most EST
                  BLASTX
Method
                   q3080412
NCBI GI
BLAST score
                   209
                   3.0e-16
E value
                   129
Match length
% identity
                   33
NCBI Description (AL022604) putative protein [Arabidopsis thaliana]
                   10316
Seq. No.
                   18486 1.R1040
Contig ID
                   jC-gmf102220148c03a1
5'-most EST
                   BLASTX
Method
                   g2289003
NCBI GI
                   414
BLAST score
                   2.0e-40
E value
                   129
Match length
% identity
                   67
                  (AC002335) membrane transporter D1 isolog [Arabidopsis
NCBI Description
                   thalianal
                   10317
Seq. No.
                   18488 1.R1040
Contig ID
                   LIB3056-004-Q1-N1-B2
5'-most EST
                   10318
Seq. No.
                   18489 1.R1040
Contig ID
                   LIB3092-043-Q1-K1-E6
5'-most EST
Method
                   BLASTX
                   q4519673
NCBI GI
BLAST score
                   1106
                   1.0e-121
E value
                   280
Match length
                   71
```

1806

NCBI Description (AB017694) WREBP-2 [Nicotiana tabacum]

% identity



```
10319
Seq. No.
                  18497 1.R1040
Contig ID
                  zzp700833272.h1
5'-most EST
                  BLASTX
Method
                  g1914685
NCBI GI
                  763
BLAST score
                  2.0e-91
E value
                  242
Match length
                  77
% identity
                  (Y12014) RAD23 protein, isoform II [Daucus carota]
NCBI Description
Seq. No.
                  10320
                  18498 1.R1040
Contig ID
                   jsh701064031.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                  g3395441
                   282
BLAST score
E value
                   5.0e-25
                   71
Match length
                   72
% identity
                  (AC004683) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10321
Seq. No.
                   18499 1.R1040
Contig ID
                   LIB3170-021-Q1-K1-C11
5'-most EST
                   BLASTN
Method
NCBI GI
                   g2564045
BLAST score
                   47
                   2.0e-17
E value
                   184
Match length
                   85
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:
NCBI Description
                   K8K14, complete sequence [Arabidopsis thaliana]
                   10322
Seq. No.
                   18507 1.R1040
Contig ID
                   LIB3170-023-Q1-K1-C5
5'-most EST
Method
                   BLASTX
                   g3892045
NCBI GI
BLAST score
                   224
E value
                   2.0e-18
Match length
                   43
% identity
                   98
                   (AC002330) putative zinc finger protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   10323
                   18509 1.R1040
Contig ID
                   fua701038279.hl
5'-most EST
                   BLASTX
Method
                   g416758
NCBI GI
                   1808
BLAST score
                   0.0e+00
E value
                   455
Match length
                   73
```

NCBI Description SERINE CARBOXYPEPTIDASE PRECURSOR >gi_166674 (M81130)



carboxypeptidase Y-like protein [Arabidopsis thaliana]
>gi_445120_prf__1908426A carboxypeptidase Y [Arabidopsis thaliana]

Seq. No. 10324 Contig ID 18510 1.R1040

5'-most EST LIB3049-007-Q1-E1-H6

Method BLASTN NCBI GI g2760167

NCBI GI g276016
BLAST score 36
E value 6.0e-11
Match length 80
% identity 86

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MCO15, complete sequence [Arabidopsis thaliana]

Seq. No. 10325

Contig ID 18510 2.R1040 5'-most EST hyd700730781.h1

Method BLASTN
NCBI GI g2760167
BLAST score 38
E value 5.0e-12
Match length 134
% identity 82

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MCO15, complete sequence [Arabidopsis thaliana]

Seq. No. 10326

Contig ID 18511 1.R1040

5'-most EST LIB3094-065-Q1-K1-E9

Method BLASTX
NCBI GI g3063701
BLAST score 544
E value 2.0e-55
Match length 144
% identity 73

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 10327

Contig ID 18511 2.R1040

5'-most EST LIB31 $\overline{7}$ 0-020-Q1-J1-F4

Method BLASTX
NCBI GI g3063701
BLAST score 384
E value 5.0e-37
Match length 98
% identity 76

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 10328

Contig ID 18518_1.R1040

5'-most EST LIB3138-033-Q1-N1-C5

Seq. No. 10329

Contig ID 18519 1.R1040

5'-most EST LIB3049-007-Q1-E1-G9



```
10330
Seq. No.
                  18520 1.R1040
Contig ID
                  jex700910091.h1
5'-most EST
                  10331
Seq. No.
                  18529 1.R1040
Contig ID
                  uC-qmropic065c11b1
5'-most EST
                  BLASTX
Method
                  q1480670
NCBI GI
BLAST score
                  503
                  1.0e-50
E value
                  151
Match length
% identity
                  (U60267) delta 1-pyrroline-5-carboxylate synthetase
NCBI Description
                   [Lycopersicon esculentum]
                  10332
Seq. No.
                  18535 1.R1040
Contig ID
                  bth700847296.h1
5'-most EST
                  BLASTX
Method
                   a3128209
NCBI GI
                   324
BLAST score
                   8.0e-30
E value
                   81
Match length
% identity
                  (AC004077) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10333
Seq. No.
                   18538 1.R1040
Contig ID
                   LIB3107-066-Q1-K1-F2
5'-most EST
                   BLASTX
Method
NCBI GI
                   q2494041
BLAST score
                   803
                   2.0e-85
E value
                   281
Match length
                   57
% identity
                  DIAMINOPIMELATE EPIMERASE >gi 1653875_dbj_BAA18785
NCBI Description
                   (D90917) diaminopimelate epimerase [Synechocystis sp.]
                   10334
Seq. No.
                   18538 2.R1040
Contig ID
                   tku700646390.hl
5'-most EST
                   BLASTX
Method
                   q2494041
NCBI GI
                   361
BLAST score
                   4.0e-34
E value
                   88
Match length
                   72
 % identity
                   DIAMINOPIMELATE EPIMERASE >gi 1653875 dbj BAA18785
NCBI Description
                   (D90917) diaminopimelate epimerase [Synechocystis sp.]
                   10335
 Seq. No.
                   18538 3.R1040
 Contig ID
                   dpv701102513.hl
 5'-most EST
```

1809

10336

Seq. No.

BLAST score

E value

304

3.0e-41



```
18540 1.R1040
Contig ID
                  LIB3051-079-Q1-K1-E9
5'-most EST
Method
                  BLASTX
                  g2464901
NCBI GI
BLAST score
                  205
                  4.0e-16
E value
Match length
                  88
% identity
                  48
                  (Z99708) putative protein [Arabidopsis thaliana]
NCBI Description
                  10337
Seq. No.
                  18542 1.R1040
Contig ID
                  LIB31\overline{0}9-014-Q1-K1-E1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4191791
BLAST score
                  1202
                  1.0e-132
E value
Match length
                  324
% identity
                   69
                   (AC005917) putative sf21 {Helianthus annuus} protein
NCBI Description
                   [Arabidopsis thaliana]
                  10338
Seq. No.
                  18542 2.R1040
Contig ID
                  uC-gmronoir040h10b1
5'-most EST
Method
                  BLASTX
                   g4191791
NCBI GI
BLAST score
                   499
E value
                   1.0e-50
                   135
Match length
                   69
% identity
                   (AC005917) putative sf21 {Helianthus annuus} protein
NCBI Description
                   [Arabidopsis thaliana]
                   10339
Seq. No.
                   18542 4.R1040
Contig ID
                   jC-gmst02400074g04d1
5'-most EST
                   10340
Seq. No.
                   18542 5.R1040
Contig ID
                   seb700651041.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q4191791
BLAST score
                   142
E value
                   7.0e-09
Match length
                   34
                   76
% identity
                  (AC005917) putative sf21 {Helianthus annuus} protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                   10341
                   18542 6.R1040
Contig ID
5'-most EST
                   uC-qmrominsoy087f02b1
Method
                   BLASTX
NCBI GI
                   q4191791
```

1810



Match length 120 % identity 71

(AC005917) putative sf21 {Helianthus annuus} protein NCBI Description

[Arabidopsis thaliana]

Seq. No. 10342

18549 1.R1040 Contig ID 5'-most EST g5058258

10343 Seq. No.

Contig ID 18549 2.R1040

5'-most EST jC-gmst02400061a04a1

10344 Seq. No.

18549 3.R1040 Contig ID

5'-most EST jC-gmro02910075f07a1

Seq. No. 10345

18549 4.R1040 Contig ID 5'-most EST leu701149641.h1

BLASTX Method NCBI GI a3095111 BLAST score 193 2.0e-14 E value 144 Match length 33 % identity

(AF051894) 15 kDa selenoprotein [Homo sapiens] NCBI Description

Seq. No.

18560 1.R1040 Contig ID

uC-gmrominsoy112f04b1 5'-most EST

10346

Method BLASTX q3757515 NCBI GI 168 BLAST score 2.0e-11 E value 82 Match length % identity 41

(AC005167) hypothetical protein [Arabidopsis thaliana] NCBI Description

10347 Seq. No.

18566 1.R1040 Contig ID

LIB3049-007-Q1-E1-A8 5'-most EST

10348 Seq. No.

18567 1.R1040 Contig ID

5'-most EST LIB3170-023-Q1-J1-C12

10349 Seq. No.

18569 1.R1040 Contig ID

5'-most EST LIB3093-013-Q1-K1-F1

Seq. No. 10350

18572 1.R1040 Contig ID

5'-most EST LIB3050-023-Q1-K1-A8

BLASTX Method NCBI GI g3237190 BLAST score 237



```
2.0e-19
E value
Match length
                   92
% identity
                  (AB014760) cystein proteinase inhibitor [Cucumis sativus]
NCBI Description
Seq. No.
                   10351
                   18578 1.R1040
Contig ID
                  LIB3109-006-Q1-K1-B8
5'-most EST
                   10352
Seq. No.
Contig ID
                   18583 1.R1040
                   jC-gmf102220068h02d1
5'-most EST
                   10353
Seq. No.
                   18585 1.R1040
Contig ID
5'-most EST
                   jC-gmf102220102f03a1
                   BLASTX
Method
NCBI GI
                   a542020
                   493
BLAST score
                   1.0e-49
E value
                   155
Match length
                   62
% identity
                   sucrose transport protein - castor bean
NCBI Description
                   >gi 468562_emb_CAA83436_ (Z31561) sucrose carrier [Ricinus
                   communis]
                   10354
Seq. No.
                   18586 1.R1040
Contig ID
                   q4314\overline{0}18
5'-most EST
Method
                   BLASTX
                   g3702343
NCBI GI
BLAST score
                   1625
E value
                   0.0e + 00
Match length
                   468
% identity
                   76
                   (ACO05397) putative homeotic gene regulator [Arabidopsis
NCBI Description
                   thaliana]
                   10355
Seq. No.
                   18590 1.R1040
Contig ID
5'-most EST
                   ncj700987760.hl
Seq. No.
                   10356
                   18590 2.R1040
Contig ID
                   kl1701214752.h1
5'-most EST
                   10357
Seq. No.
                   18593 1.R1040
Contig ID
                   LIB3049-006-Q1-E1-H3
 5'-most EST
                   BLASTX
Method
                   g2244765
NCBI GI
BLAST score
                   285
                   2.0e-25
E value
                   94
Match length
```

NCBI Description (Z97335) hypothetical protein [Arabidopsis thaliana]

54

% identity



18595 1.R1040 Contig ID 5'-most EST kl1701213810.h1

Method BLASTX g3582436 NCBI GI 1153 BLAST score 1.0e-148 E value 345 Match length 75 % identity

(AB017502) beta-D-glucan exohydrolase [Nicotiana tabacum] NCBI Description

10359 Seq. No.

18598 1.R1040 Contig ID

5'-most EST LIB3049-006-Q1-E1-G12

Seq. No. 10360

18610 1.R1040 Contig ID

LIB3049-006-Q1-E1-F6 5'-most EST

10361 Seq. No.

18611 1.R1040 Contig ID

LIB3049-006-Q1-E1-F7 5'-most EST

10362 Seq. No.

18613 1.R1040 Contig ID

LIB3049-006-Q1-E1-G1 5'-most EST

Seq. No. 10363

Contig ID 18615 1.R1040

jC-qmro02800022f01a1 5'-most EST

BLASTX Method q4415939 NCBI GI 533 BLAST score 6.0e-54 E value 250 Match length % identity 55

(AC006418) putative zinc finger protein [Arabidopsis NCBI Description

thaliana]

10364 Seq. No.

18615 3.R1040 Contig ID trc700565847.h1 5'-most EST

BLASTX Method g4415939 NCBI GI BLAST score 348 E value 8.0e-33 71 Match length

% identity (AC006418) putative zinc finger protein [Arabidopsis NCBI Description

thaliana]

10365 Seq. No.

18615 4.R1040 Contig ID uC-gmropic068g10b1 5'-most EST

86

BLASTX Method NCBI GI g4415939 BLAST score 380

1813



E value 1.0e-36
Match length 118
% identity 67
NCBI Description (AC006418) putative zinc finger protein [Arabidopsis thaliana]
Seq. No. 10366

Contig ID 18618_1.R1040
5'-most EST vzy700753764.h1
Method BLASTX
NCBI GI g3600052
BLAST score 635
E value 6.0e-66
Match length 311

% identity 44

NCBI Description (AF080120) contains similarity to glycosyl hydrolases family 9 (Pfam: glycosyl hydro5.hmm, score: 88.03)

[Arabidopsis thaliana]

Seq. No. 10367

Contig ID 18620_1.R1040

5'-most EST LIB3049-006-Q1-E1-D3

Seq. No. 10368

Contig ID 18622_1.R1040 5'-most EST leu701157390.h1

Method BLASTX
NCBI GI g1708420
BLAST score 271
E value 8.0e-24
Match length 123
% identity 43

NCBI Description ISOFLAVONE REDUCTASE HOMOLOG P3 >gi_1361992_pir__S57613 isoflavonoid reductase homolog - Arabidopsis thaliana

isoflavonoid reductase homolog - Arabidopsis thalland >gi_886432_emb_CAA89859_ (Z49777) isoflavonoid reductase

homologue [Arabidopsis thaliana]

Seq. No. 10369

Contig ID 18623_1.R1040

5'-most EST LIB3139-117-P1-N1-G9

Method BLASTX
NCBI GI g2832623
BLAST score 168
E value 1.0e-11
Match length 61
% identity 49

NCBI Description (AL021711) protein kinase - like protein [Arabidopsis

thaliana]

Seq. No. 10370

Contig ID 18624 1.R1040 5'-most EST zhf700957067.h1

Method BLASTN
NCBI GI g2792154
BLAST score 593
E value 0.0e+00
Match length 981



% identity 90
NCBI Description Sesbania rostrata mRNA for putative chalcone reductase

Seq. No. 10371

Contig ID 18631_1.R1040

5'-most EST LIB3049-006-Q1-E1-E3

Method BLASTN
NCBI GI g169157
BLAST score 136
E value 2.0e-70
Match length 376
% identity 87

NCBI Description Pisum sativum serine hydroxymethyltransferase mRNA,

complete cds

Seq. No. 10372

Contig ID 18633_1.R1040

5'-most EST LIB3049-006-Q1-E1-D10

Seq. No. 10373

Contig ID 18635_1.R1040 5'-most EST zzp700832155.h1

Method BLASTX
NCBI GI g2213626
BLAST score 193
E value 2.0e-14
Match length 156
% identity 38

NCBI Description (AC000103) F21J9.18 [Arabidopsis thaliana]

Seq. No. 10374

Contig ID 18635_2.R1040 5'-most EST uaw700660817.h1

Seq. No. 10375

Contig ID 18645_1.R1040

5'-most EST LIB3049-006-Q1-E1-C5

Method BLASTX
NCBI GI g2880043
BLAST score 734
E value 1.0e-77
Match length 300
% identity 46

NCBI Description (AC002340) putative 3-hydroxyisobutyryl-coenzyme A

hydrolase [Arabidopsis thaliana]

Seq. No. 10376

Contig ID 18647_1.R1040

5'-most EST LIB3049-006-Q1-E1-C7

Seq. No. 10377

Contig ID 18648 1.R1040

5'-most EST LIB3092-061-Q1-K1-C1

Seq. No. 10378

Contig ID 18651 1.R1040

5'-most EST LIB3109-050-Q1-K1-G2



Contig ID 18652_1.R1040

5'-most EST LIB3170-023-Q1-K1-B9

Seq. No. 10380

Contig ID 18656_1.R1040

5'-most EST LIB3049-006-Q1-E1-A3

Method BLASTX
NCBI GI g3548801
BLAST score 346
E value 3.0e-32
Match length 192

Match length 192 % identity 36

NCBI Description (AC005313) putative transmembrane protein [Arabidopsis thaliana] >gi 4335768 gb_AAD17445_ (AC006284) putative

integral membrane protein [Arabidopsis thaliana]

Seq. No. 10381

Contig ID 18656 2.R1040 5'-most EST leu701147073.h1

Method BLASTX
NCBI GI g3461828
BLAST score 157
E value 2.0e-10
Match length 81

% identity 38
NCBI Description (AC004138) unknown protein [Arabidopsis thaliana]

Seq. No. 10382

Contig ID 18657_1.R1040 5'-most EST uaw700665034.h1

Method BLASTX
NCBI GI 94468048
BLAST score 336
E value 2.0e-31
Match length 152
% identity 46

NCBI Description (X78703) catechol O-methyltransferase [Vanilla planifolia]

Seq. No. 10383

Contig ID 18659_1.R1040

5'-most EST LIB3049-006-Q1-E1-A6

Seq. No.

10384

Contig ID 18663_1.R1040

5'-most EST LIB3106-095-Q1-K1-F10

Seq. No. 10385

Contig ID 18669_1.R1040

5'-most EST uC-gmrominsoy031f01b1

Seq. No.

10386

Contig ID 18669_2.R1040

5'-most EST LIB3106-013-Q1-K1-E6

Seq. No. 10387



Contig ID 18670_1.R1040

5'-most EST LIB3049-005-Q1-E1-G7

Method BLASTX
NCBI GI g1086147
BLAST score 450
E value 1.0e-44
Match length 131
% identity 61

NCBI Description protein S2 - Phalaris coerulescens >gi_556833_emb_CAA57520_

(X81992) S2 [Phalaris coerulescens]

Seq. No. 10388

Contig ID 18670 2.R1040

5'-most EST LIB3107-015-Q1-K1-E10

Seq. No. 10389

Contig ID 18670_3.R1040 5'-most EST vzy700751450.h1

Seq. No. 10390

Contig ID 18678 1.R1040

5'-most EST LIB3049-005-Q1-E1-H4

Method BLASTX
NCBI GI g2498329
BLAST score 593
E value 1.0e-61
Match length 127
% identity 91

NCBI Description PATTERN FORMATION PROTEIN EMB30 >gi_2129665_pir__\$65571

pattern-formation protein GNOM - Arabidopsis thaliana

>gi_1209633 (U36433) GNOM gene product [Arabidopsis

thaliana] >gi_1335997 (U56140) similar to the Saccharomyces cerevisiae Sec7 protein, GenBank Accession Number J03918 [Arabidopsis thaliana] >gi_1335999 (U56141) similar to the Saccharomyces cerevisiae Sec7 protein, GenBank Accession

Number J03918 [Arabidopsis thaliana]

Seq. No. 10391

Contig ID 18680 1.R1040

5'-most EST LIB3094-043-Q1-K1-C9

Method BLASTX
NCBI GI g4544399
BLAST score 745
E value 3.0e-79
Match length 177
% identity 81

NCBI Description (AC007047) putative beta-ketoacyl-CoA synthase [Arabidopsis

thaliana]

Seq. No. 10392

Contig ID 18683 1.R1040

5'-most EST LIB3049-005-Q1-E1-H9

Seq. No. 10393

Contig ID 18686 1.R1040

5'-most EST LIB3049-005-Q1-E1-F10

Method BLASTX



```
q629858
NCBI GI
                  533
BLAST score
                  1.0e-54
E value
                  127
Match length
% identity
NCBI Description protein kinase C inhibitor - maize
                  10394
Seq. No.
                   18687 1.R1040
Contig ID
                  LIB3049-005-Q1-E1-F11
5'-most EST
                   10395
Seq. No.
                   18688 1.R1040
Contig ID
5'-most EST
                   leu701151019.h1
                   BLASTX
Method
NCBI GI
                   q2129675
BLAST score
                   638
                   1.0e-66
E value
Match length
                   192
                   69
% identity
NCBI Description
                   >gi_972938 (U19382) putative chlorophyll synthetase
                   chlorophyll synthetase [Arabidopsis thaliana]
                   10396
```

probable chlorophyll synthetase G4 - Arabidopsis thaliana

[Arabidopsis thaliana] >gi_3068709 (AF049236) putative

Seq. No.

18689 1.R1040 Contig ID

LIB3049-005-Q1-E1-F2 5'-most EST

10397 Seq. No.

18689 2.R1040 Contig ID asn701132169.hl 5'-most EST

10398 Seq. No.

18691 1.R1040 Contig ID zsg701127305.h1 5'-most EST

Method BLASTX g3927830 NCBI GI 506 BLAST score 3.0e-57 E value Match length 200 % identity 28

(AC005727) hypothetical protein [Arabidopsis thaliana] NCBI Description

Seq. No. 10399

18694 1.R1040 Contig ID

uC-qmflminsoy091e09b1 5'-most EST

Method BLASTN q3941288 NCBI GI 376 BLAST score 0.0e+00E value 828 Match length % identity 86

Pisum sativum similarity to SCAMP37 (psam2) mRNA, complete NCBI Description

10400 Seq. No.



Contig ID 18694_2.R1040
5'-most EST LIB3051-047-Q1-K1-H9
Method BLASTN
NCBI GI g3941288
BLAST score 92
E value 5.0e-44

NCBI Description Pisum sativum similarity to SCAMP37 (psam2) mRNA, complete

cds

249

87

Seq. No. 10401

Match length

% identity

Contig ID 18694_3.R1040

5'-most EST jC-gmro02910054h10a1

Method BLASTN
NCBI GI g3941288
BLAST score 48
E value 8.0e-18
Match length 145
% identity 88

NCBI Description Pisum sativum similarity to SCAMP37 (psam2) mRNA, complete

cds

Seq. No. 10402

Contig ID 18694 4.R1040

5'-most EST jC-gmro02910054h10d1

Method BLASTN
NCBI GI g3941288
BLAST score 215
E value 1.0e-117
Match length 379
% identity 89

NCBI Description Pisum sativum similarity to SCAMP37 (psam2) mRNA, complete

cds

Seq. No. 10403

Contig ID 18702_1.R1040 5'-most EST kl1701202317.h1

Method BLASTX
NCBI GI 94490316
BLAST score 554
E value 4.0e-57
Match length 229
% identity 53

NCBI Description (AL035678) nucellin-like protein [Arabidopsis thaliana]

Seq. No. 10404

Contig ID 18702_2.R1040

5'-most EST LIB3170-047-Q1-J1-B3

Method BLASTX
NCBI GI g4490316
BLAST score 247
E value 2.0e-33
Match length 131
% identity 53

NCBI Description (AL035678) nucellin-like protein [Arabidopsis thaliana]



```
10405
Seq. No.
                   18704 1.R1040
Contig ID
                  LIB3050-020-Q1-K1-G1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3860258
BLAST score
                   408
                   9.0e-40
E value
Match length
                   118
% identity
                   (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10406
                   18704 2.R1040
Contig ID
                   LIB31\overline{7}0-021-Q1-K1-B7
5'-most EST
                   BLASTX
Method
NCBI GI
                   q3860258
BLAST score
                   162
E value
                   3.0e-11
                   52
Match length
                   60
% identity
                   (AC005824) unknown protein [Arabidopsis thaliana]
NCBI Description
                   10407
Seq. No.
                   18705 1.R1040
Contig ID
                   uC-gmflminsoy080h11b1
5'-most EST
                   BLASTX
Method
                   g2827528
NCBI GI
BLAST score
                   436
E value
                   9.0e-43
Match length
                   216
% identity
                   (AL021633) predicted protein [Arabidopsis thaliana]
NCBI Description
                   10408
Seq. No.
                   18707 1.R1040
Contig ID
                   kmv700741951.hl
5'-most EST
                   10409
Seq. No.
                   18707 2.R1040
Contig ID
5'-most EST
                   zzp700836141.hl
                   10410
Seq. No.
                   18710 1.R1040
Contig ID
                   LIB3170-022-Q1-K1-B9
5'-most EST
                   BLASTX
Method
                   g3269284
NCBI GI
                   395
BLAST score
                   2.0e-38
E value
                   128
Match length
                   62
% identity
                   (AL030978) histone H2A- like protein [Arabidopsis thaliana]
NCBI Description
                   10411
Seq. No.
```

18729 1.R1040 Contig ID

LIB3092-015-Q1-K1-G11 5'-most EST

Method BLASTX q4105798 NCBI GI



BLAST score 931
E value 1.0e-100
Match length 249
% identity 57
NCBI Description (AF049930) PGP237-11 [Petunia x hybrida]

Seq. No. 10412

Contig ID 18729_2.R1040

5'-most EST LIB3109-007-Q1-K1-C1

Method BLASTX
NCBI GI g4105798
BLAST score 207
E value 3.0e-16
Match length 65
% identity 54

NCBI Description (AF049930) PGP237-11 [Petunia x hybrida]

Seq. No. 10413

Contig ID 18732_1.R1040

5'-most EST LIB3049-005-Q1-E1-C6

Method BLASTX
NCBI GI g1352881
BLAST score 146
E value 7.0e-09
Match length 110
% identity 28

NCBI Description HYPOTHETICAL 15.7 KD PROTEIN IN NUP85-SSC1 INTERGENIC

REGION >gi_1078266_pir__S57063 probable membrane protein

YJR044c - yeast (Saccharomyces cerevisiae) >qi 1015699 emb CAA89572 (Z49544) ORF YJR044c

[Saccharomyces cerevisiae] >gi_1197072 (L36344) ORF;

putative [Saccharomyces cerevisiae]

Seq. No. 10414

Contig ID 18732_2.R1040

5'-most EST LIB3056-014-Q1-N1-D1

Seq. No. 10415

Contig ID 18738_1.R1040

5'-most EST LIB3049-004-Q1-E1-H4

Seq. No. 10416

Contig ID 18741_1.R1040 5'-most EST zpv700761129.h1

Method BLASTX
NCBI GI g3451411
BLAST score 530
E value 4.0e-54
Match length 145
% identity 70

NCBI Description (Z98761) seryl-tRNA synthetase [Helianthus annuus]

Seq. No. 10417

Contig ID 18746_1.R1040

5'-most EST LIB3049-005-Q1-E1-A5

Method BLASTX NCBI GI g1843628



465 BLAST score 2.0e-46 E value Match length 127 72 % identity (U88061) SNF5 homolog BSH [Arabidopsis thaliana] NCBI Description 10418 Seq. No. 18749 1.R1040 Contig ID LIB3170-021-Q1-K1-B5 5'-most EST Seq. No. 10419 18750 1.R1040 Contig ID $LIB30\overline{4}9-005-Q1-E1-B1$ 5'-most EST BLASTX Method q3970680 NCBI GI BLAST score 275 E value 3.0e-24Match length 108 % identity (AL034388) 67A9.b [Drosophila melanogaster] NCBI Description 10420 Seq. No. 18750 2.R1040 Contig ID pmv700889336.h1 5'-most EST BLASTX Method g3970680 NCBI GI BLAST score 105 1.0e-11 E value 99 Match length % identity 46 (AL034388) 67A9.b [Drosophila melanogaster] NCBI Description 10421 Seq. No. 18750 3.R1040 Contia ID uC-gmronoir064h03b1 5'-most EST Seq. No. 10422 18750 4.R1040 Contig ID jex700906670.hl 5'-most EST Seq. No. 10423 18751 1.R1040 Contig ID sat701011064.hl 5'-most EST BLASTX Method g3395938 NCBI GI 781 BLAST score 4.0e-83 E value 251 Match length 50 % identity (AF076924) polypyrimidine tract-binding protein homolog NCBI Description [Arabidopsis thaliana]

10424 Seq. No.

18753 1.R1040 Contig ID

LIB3170-022-Q1-K1-B2 5'-most EST

BLASTX Method NCBI GI g3269287

1822



```
174
BLAST score
                   1.0e-12
E value
                   58
Match length
                   62
% identity
NCBI Description (AL030978) GH3 like protein [Arabidopsis thaliana]
Seq. No.
                   10425
                   18754 1.R1040
Contig ID
                   uC-gmronoir063d11b1
5'-most EST
                   10426
Seq. No.
                   18760 1.R1040
Contig ID
                   trc700564090.hl
5'-most EST
                   10427
Seq. No.
                   18760 2.R1040
Contig ID
                   LIB3051-051-Q1-K1-H1
5'-most EST
                   10428
Seq. No.
                   18768 1.R1040
Contig ID
                   LIB3049-004-Q1-E1-H2
5'-most EST
                   BLASTX
Method
                   g3687240
NCBI GI
                   227
BLAST score
                   7.0e-19
E value
                   100
Match length
                   42
% identity
NCBI Description (AC005169) extensin-like protein [Arabidopsis thaliana]
                   10429
Seq. No.
                   18779 1.R1040
Contig ID
                   zzp700831511.hl
5'-most EST
                   BLASTX
Method
                   q2661179
NCBI GI
                   800
BLAST score
                   1.0e-85
E value
                   215
Match length
% identity
                   (U80984) AtZW10 [Arabidopsis thaliana]
NCBI Description
                   10430
 Seq. No.
                   18780 1.R1040
Contig ID
                   LIB31\overline{3}9-034-P1-N1-E12
 5'-most EST
Method
                   BLASTX
                   g2829918
NCBI GI
BLAST score
                   153
                   4.0e-11
E value
Match length
                   59
 % identity
                   71
                   (AC002291) similar to "tub" protein gp_U82468_2072162
 NCBI Description
                    [Arabidopsis thaliana]
```

Seq. No. 10431 Contig ID 18780_2.R1040

5'-most EST LIB3049-004-Q1-E1-F10

Method BLASTX NCBI GI g2829918



```
302
BLAST score
                   2.0e-27
E value
Match length
                   85
                   73
% identity
                   (AC002291) similar to "tub" protein gp_U82468_2072162
NCBI Description
                   [Arabidopsis thaliana]
                   10432
Seq. No.
                   18780 5.R1040
Contig ID
5'-most EST
                   dpv701099903.h1
                   10433
Seq. No.
                   18783 1.R1040
Contig ID
5'-most EST
                   LIB3049-004-Q1-E1-F2
Seq. No.
                   10434
Contig ID
                   18788 1.R1040
5'-most EST
                   xpa700792484.h1
                   BLASTX
Method
                   q2688824
NCBI GI
                   257
BLAST score
E value
                   8.0e-22
                   126
Match length
                   52
% identity
                   (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                   armeniaca]
                   10435
Seq. No.
                   18788 2.R1040
Contig ID
5'-most EST
                   q5605911
                   BLASTX
Method
NCBI GI
                   g2688824
BLAST score
                   243
                   3.0e-20
E value
Match length
                   140
% identity
                   45
                   (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                   armeniaca]
                   10436
Seq. No.
Contig ID
                   18788 4.R1040
                   asn701141515.hl
5'-most EST
Method
                   BLASTX
                   q2688824
NCBI GI
                   169
BLAST score
                   4.0e-12
E value
                   40
Match length
% identity
                   78
                   (U93273) putative auxin-repressed protein [Prunus
NCBI Description
                   armeniaca]
                   10437
Seq. No.
                   18789 1.R1040
Contig ID
```

LIB3050-003-Q1-E1-D6 5'-most EST

10438 Seq. No.

Contig ID 18791 1.R1040



```
5'-most EST
                  LIB3109-035-Q1-K1-C2
                  BLASTX
Method
                  g2346986
NCBI GI
                  397
BLAST score
E value
                  4.0e-38
Match length
                  263
% identity
                  44
NCBI Description (AB006605) ZPT3-3 [Petunia x hybrida]
                  10439
Seq. No.
Contig ID
                  18798 1.R1040
                  LIB3093-047-Q1-K1-G10
5'-most EST
                  BLASTX
Method
NCBI GI
                  g623586
BLAST score
                  960
                  1.0e-104
E value
Match length
                  207
                  95
% identity
NCBI Description (L29273) putative [Nicotiana tabacum]
                  10440
Seq. No.
Contig ID
                  18798 2.R1040
                  LIB3049-004-Q1-E1-D4
5'-most EST
Method
                  BLASTN
                  g435428
NCBI GI
BLAST score
                  76
E value
                  1.0e-34
                  152
Match length
% identity
NCBI Description Arabidopsis thaliana GTP-binding protein mRNA
                  10441
Seq. No.
Contig ID
                  18798 3.R1040
5'-most EST
                  vwf700676648.h1
                  BLASTN
Method
NCBI GI
                  q435428
BLAST score
                  92
E value
                   4.0e-44
Match length
                  188
                  87
% identity
NCBI Description Arabidopsis thaliana GTP-binding protein mRNA
Seq. No.
                  10442
                  18799 1.R1040
Contig ID
                  LIB3049-004-Q1-E1-D5
5'-most EST
Method
                  BLASTX
                  g2500354
NCBI GI
BLAST score
                  1073
                   1.0e-117
E value
                  219
Match length
% identity
                   91
                  60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894_dbj_BAA19462_
NCBI Description
                   (AB001891) QM family protein [Solanum melongena]
```

Contig ID 18799_2.R1040

5'-most EST LIB3051-024-Q1-K1-F4



BLASTX Method g2500354 NCBI GI BLAST score 542 2.0e-55 E value Match length 117 % identity 85 60S RIBOSOMAL PROTEIN L10 (EQM) >gi 1902894 dbj_BAA19462_ NCBI Description (AB001891) QM family protein [Solanum melongena] 10444 Seq. No. 18801 1.R1040 Contig ID g5677221 5'-most EST Seq. No. 10445 Contig ID 18801 2.R1040 5'-most EST LIB3094-082-Q1-K1-D2 Method BLASTX NCBI GI q4220534 BLAST score 494 1.0e-49 E value Match length 199 % identity 51 NCBI Description (AL035356) putative protein [Arabidopsis thaliana] 10446 Seq. No. Contig ID 18814 1.R1040 $asn70\overline{1}137579.h1$ 5'-most EST Seq. No. 10447 18815 1.R1040 Contig ID $1eu70\overline{1}151423.h1$ 5'-most EST 10448 Seq. No. 18816 1.R1040 Contig ID LIB3170-023-Q1-K1-B11 5'-most EST 10449 Seq. No. 18817 1.R1040 Contig ID 5'-most EST LIB3170-023-Q1-K1-B3 10450 Seq. No. 18820 1.R1040 Contig ID LIB3049-004-Q1-E1-C2 5'-most EST BLASTX Method g2244800 NCBI GI BLAST score 326 2.0e-30 E value 106 Match length

% identity

(Z97336) carnitine racemase homolog [Arabidopsis thaliana] NCBI Description

Seq. No.

10451

Contig ID

18821 1.R1040

5'-most EST

LIB3170-024-Q1-K1-B7

Seq. No.

10452

Contig ID

18823 1.R1040



```
LIB3170-024-Q1-K1-B8
5'-most EST
Seq. No.
                  10453
Contig ID
                  18830 1.R1040
                  jC-gmf102220075b05d1
5'-most EST
                  BLASTX
Method
                  q3193296
NCBI GI
BLAST score
                  848
                  5.0e-91
E value
                  224
Match length
% identity
                  67
                  (AF069298) similar to pectinesterase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  10454
                  18831 1.R1040
Contig ID
                  LIB3049-004-Q1-E1-A10
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3953479
BLAST score
                   173
E value
                   3.0e-12
                   58
Match length
% identity
                   57
                  (AC002328) F2202.24 [Arabidopsis thaliana]
NCBI Description
                   10455
Seq. No.
                   18832 1.R1040
Contig ID
                   LIB3170-023-Q1-K1-B10
5'-most EST
Seq. No.
                   10456
                   18834 1.R1040
Contig ID
                   LIB3170-023-Q1-K1-B8
5'-most EST
                   10457
Seq. No.
                   18842 1.R1040
Contiq ID
                   LIB3049-003-Q1-E1-F9
5'-most EST
                   BLASTX
Method
                   g2462737
NCBI GI
                   309
BLAST score
                   2.0e-28
E value
Match length
                   126
                   49
% identity
                   (AC002292) similar to "Mx" GTP-binding proteins
NCBI Description
                   [Arabidopsis thaliana]
                   10458
Seq. No.
                   18844 1.R1040
Contig ID
                   gsv701047810.hl
5'-most EST
                   10459
Seq. No.
                   18844 2.R1040
Contig ID
5'-most EST
                   LIB3056-010-Q1-N1-E12
                   10460
```

Contig ID 18849 1.R1040

uC-gmrominsoy153d04b1 5'-most EST

Method BLASTX q3004549 NCBI GI



```
BLAST score
                  358
                  2.0e-33
E value
                  177
Match length
% identity
                  45
                  (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                  >qi 4185152 (AC005724) unknown protein [Arabidopsis
                  thaliana]
                  10461
Seq. No.
                  18849 2.R1040
Contig ID
5'-most EST
                  LIB3049-003-Q1-E1-G4
Method
                  BLASTX
                  g3004549
NCBI GI
                  155
BLAST score
                   4.0e-13
E value
Match length
                   95
% identity
                   (AC003673) unknown protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4185152 (AC005724) unknown protein [Arabidopsis
                   thaliana]
                   10462
Seq. No.
                   18851 1.R1040
Contig ID
                   LIB3049-003-Q1-E1-G6
5'-most EST
                   BLASTX
Method
                   q4105697
NCBI GI
BLAST score
                   984
                   1.0e-107
E value
                   233
Match length
% identity
                   77
                  (AF049870) thaumatin-like protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10463
                   18851 3.R1040
Contig ID
                   uC-gmrominsoy106b10b1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4105697
BLAST score
                   369
                   3.0e-35
E value
Match length
                   86
% identity
                   74
                  (AF049870) thaumatin-like protein [Arabidopsis thaliana]
NCBI Description
                   10464
Seq. No.
                   18852 1.R1040
Contig ID
                   LIB3170-024-Q1-K1-A1
5'-most EST
Method
                   BLASTX
                   g3036840
NCBI GI
BLAST score
                   216
                   1.0e-17
E value
Match length
                   87
% identity
                   48
                   (AJ222967) cystinosin [Homo sapiens]
NCBI Description
                   >gi_3036851_emb_CAA75882_ (Y15924) cystinosin [Homo
```

sapiens]



```
18865 1.R1040
Contig ID
                  LIB3049-003-Q1-E1-E8
5'-most EST
                  BLASTX
Method
                  g3228517
NCBI GI
BLAST score
                  263
                  7.0e-23
E value
Match length
                  97
% identity
                  63
NCBI Description (AF007788) ETTIN [Arabidopsis thaliana]
                  10466
Seq. No.
                  18867 1.R1040
Contig ID
                  rca701002296.h1
5'-most EST
                  BLASTX
Method
                  g3080393
NCBI GI
                   963
BLAST score
                   1.0e-104
E value
                   290
Match length
% identity
                   63
                   (AL022603) NADH dehydrogenase like protein [Arabidopsis
NCBI Description
                   thaliana]
                   10467
Seq. No.
                   18867 2.R1040
Contig ID
                   LIB3138-050-Q1-N1-A6
5'-most EST
                   BLASTX
Method
                   g3080393
NCBI GI
                   289
BLAST score
                   5.0e-26
E value
                   86
Match length
% identity
                   (AL022603) NADH dehydrogenase like protein [Arabidopsis
NCBI Description
                   thaliana]
                   10468
Seq. No.
                   18875 1.R1040
Contig ID
                   uC-gmronoir070h12b1
5'-most EST
Method
                   BLASTX
                   q3927825
NCBI GI
                   1660
BLAST score
                   0.0e + 00
E value
Match length
                   347
% identity
                   (AC005727) putative dTDP-glucose 4-6-dehydratase
NCBI Description
                   [Arabidopsis thaliana]
                   10469
Seq. No.
Contig ID
                   18875 2.R1040
                   LIB3139-050-P1-N1-F9
5'-most EST
Method
                   BLASTX
                   g2351580
NCBI GI
                   583
```

Method BLASTX
NCBI GI g2351580
BLAST score 583
E value 3.0e-60
Match length 132
% identity 85

NCBI Description (U82433) thymidine diphospho-glucose 4-6-dehydratase

homolog [Prunus armeniaca]



```
10470
Seq. No.
Contig ID
                  18875 3.R1040
5'-most EST
                  LIB3139-040-P1-N1-F6
Method
                  BLASTX
NCBI GI
                  q2351580
BLAST score
                  550
E value
                  3.0e-56
                  125
Match length
                  85
% identity
                  (U82433) thymidine diphospho-glucose 4-6-dehydratase
NCBI Description
                  homolog [Prunus armeniaca]
                  10471
Seq. No.
                  18875 4.R1040
Contig ID
5'-most EST
                  LIB3049-003-Q1-E1-C5
Method
                  BLASTX
NCBI GI
                   g3927825
BLAST score
                   403
                   2.0e-39
E value
                   81
Match length
% identity
                   93
                   (AC005727) putative dTDP-glucose 4-6-dehydratase
NCBI Description
                   [Arabidopsis thaliana]
                   10472
Seq. No.
Contig ID
                   18875 5.R1040
5'-most EST
                   ncj700984373.hl
Seq. No.
                   10473
Contig ID
                   18875 6.R1040
5'-most EST
                   asn701131968.hl
Method
                   BLASTX
                   g3927825
NCBI GI
BLAST score
                   451
                   7.0e-45
E value
Match length
                   99
% identity
                   88
                  (AC005727) putative dTDP-glucose 4-6-dehydratase
NCBI Description
                   [Arabidopsis thaliana]
                   10474
Seq. No.
                   18884 1.R1040
Contig ID
                   LIB3049-003-Q1-E1-D2
5'-most EST
Method
                   BLASTX
                   g3445201
NCBI GI
BLAST score
                   268
                   5.0e-23
E value
                   224
Match length
% identity
                   46
                  (AC004786) unknown protein [Arabidopsis thaliana]
NCBI Description
```

18885 1.R1040 Contig ID 5'-most EST LIB3170-022-Q1-K1-A8

BLASTX Method NCBI GI g3176687



BLAST score 613 1.0e-63 E value 160 Match length % identity 74 (AC003671) Strong similarity to trehalose-6-phosphate NCBI Description synthase homolog from A. thaliana chromosome 4 contig gb Z97344. ESTs gb H37594, gb R65023, gb_H37578 and qb R64855 come from this gene. [Arabidopsis thaliana] Seq. No. 10476 18893 1.R1040 Contig ID fua701040501.h1 5'-most EST Method BLASTX g3298536 NCBI GI BLAST score 259 E value 2.0e-22 Match length 106 % identity 60 (AC004681) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 10477 Contig ID 18896 1.R1040 fde700871557.hl 5'-most EST Method BLASTX NCBI GI g3702339 BLAST score 990 78.00 E value 1.0e-107 Match length 321 % identity (AC005397) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 10478 18896 2.R1040 Contig ID zzp700830352.h1 5'-most EST BLASTX Method NCBI GI g3702339 153 BLAST score 5.0e-10 E value Match length 51 61 % identity (AC005397) unknown protein [Arabidopsis thaliana] NCBI Description 10479 Seq. No. 18898 1.R1040 Contig ID 5'-most EST LIB3106-011-Q1-K1-F3 BLASTN Method NCBI GI g3821780 BLAST score 36 2.0e-10 E value Match length 48 67 % identity NCBI Description Xenopus laevis cDNA clone 27A6-1

Seq. No. 10480

Contig ID 18899_1.R1040

5'-most EST LIB3049-003-Q1-E1-B5

Method BLASTX

1831



NCBI GI g4469014 BLAST score 194 1.0e-14 E value 51 Match length % identity NCBI Description (AL035602) putative protein [Arabidopsis thaliana] 10481 Seq. No. 18901 1.R1040 Contig ID 5'-most EST LIB3170-023-Q1-J1-A7 10482 Seq. No. 18903 1.R1040 Contig ID

5'-most EST fua701039094.h1
Method BLASTX
NCBI GI g3157942
BLAST score 264
E value 6.0e-38

BLAST score 264
E value 6.0eMatch length 231
% identity 45

NCBI Description (AC002131) Similar to style development-specific protein 9612 precursor gb_X55193 and pectate lyase P59 precursor

gb X15499 from Lycopersicon esculentum. [Arabidopsis

thaliana]

Seq. No. 10483

Contig ID 18903_2.R1040

5'-most EST uC-gmrominsoy303a12b1

Seq. No. 10484

Contig ID 18905 1.R1040

5'-most EST LIB3049-003-Q1-E1-C10

Seq. No. 10485

Contig ID 18906 1.R1040

5'-most EST jC-gmro02910071h05a1

Seq. No. 10486

Contig ID 18906 2.R1040

5'-most EST jC-gmle01810055c08a1

Method BLASTX
NCBI GI g1871181
BLAST score 288
E value 2.0e-25
Match length 170
% identity 43

NCBI Description (U90439) ring zinc finger protein isolog [Arabidopsis

thaliana]

Seq. No. 10487

Contig ID 18906 4.R1040

5'-most EST LIB3052-012-Q1-N1-A12

Seq. No. 10488

Contig ID 18910 1.R1040

5'-most EST LIB3107-017-Q1-K1-H9



Contig ID 18915_1.R1040

5'-most EST LIB3049-002-Q1-E1-H7

Seq. No. 10490

Contig ID 18923 1.R1040

5'-most EST LIB3170-022-Q1-K1-A2

Method BLASTX
NCBI GI g1418321
BLAST score 140
E value 1.0e-08
Match length 35
% identity 40

NCBI Description (X98669) C2H2 zinc finger protein [Arabidopsis thaliana]

>gi 2317903 (U89959) C2H2 zinc finger protein [Arabidopsis

thaliana]

Seq. No. 10491

Contig ID 18926 1.R1040

5'-most EST LIB3049-003-Q1-E1-A7

Seq. No. 10492

Contig ID 18932 1.R1040

5'-most EST LIB3049-002-Q1-E1-G1

Seq. No. 10493

Contig ID 18933_1.R1040 5'-most EST seb700653121.h1

Seq. No. 10494

Contig ID 18935 1.R1040

5'-most EST LIB $30\overline{5}1-002-Q1-E1-C8$

Method BLASTX
NCBI GI g3193316
BLAST score 454
E value 2.0e-57
Match length 160
% identity 72

NCBI Description (AF069299) contains similarity to nucleotide sugar

epimerases [Arabidopsis thaliana]

Seq. No. 10495

Contig ID 18935_2.R1040 5'-most EST kl1701209745.h1

Seq. No. 10496

Contig ID 18938 1.R1040

5'-most EST LIB3138-038-Q1-N1-G5

Seq. No. 10497

Contig ID 18938 2.R1040 5'-most EST txt700732831.h1

Seq. No. 10498

Contig ID 18939_2.R1040

5'-most EST LIB3049-002-Q1-E1-G5

Method BLASTN



NCBI GI q493019 BLAST score 38 6.0e-12 E value 92 Match length % identity Glycine max delta-aminolevulinic acid dehydratase (Alad) NCBI Description mRNA, complete cds 10499 Seq. No. 18941 1.R1040 Contig ID 5'-most EST LIB3139-122-P1-N1-E6 BLASTX Method g4204314 NCBI GI 284 BLAST score 4.0e-25 E value Match length 176 % identity 38 NCBI Description (AC003027) Hypothetical protein [Arabidopsis thaliana] 10500 Seq. No. 18949 1.R1040 Contig ID 5'-most EST LIB3170-024-Q1-K1-A7 BLASTX Method g1684913 NCBI GI 455 BLAST score 1.0e-45 E value Match length 119 % identity 76 NCBI Description (U77888) receptor-like protein kinase [Ipomoea nil] 10501 Seq. No. Contig ID 18952 1.R1040 5'-most EST LIB3049-002-Q1-E1-E5 Seq. No. 10502 Contig ID 18954 1.R1040 5'-most EST LIB3049-002-Q1-E1-E7 Seq. No. 10503 Contig ID 18956 1.R1040 5'-most EST zhf700957433.hl Method BLASTX NCBI GI g2598575 BLAST score 493 3.0e-49 E value Match length 331 % identity 34 (Y15293) MtN21 [Medicago truncatula] NCBI Description 10504 Seq. No. 18956 2.R1040 Contig ID 5'-most EST jC-gmst02400025e02a1

Seq. No. 10505

Contig ID 18956_3.R1040 5'-most EST g4291204 Method BLASTX



NCBI GI g3355480 262 BLAST score 2.0e-22 E value Match length 118 42 % identity (AC004218) Medicago nodulin N21-like protein [Arabidopsis NCBI Description thaliana] 10506 Seq. No. 18956 4.R1040 Contig ID LIB3107-067-Q1-K1-H1 5'-most EST Method BLASTX g2598575 NCBI GI 171 BLAST score 1.0e-12 E value 112 Match length 36 % identity (Y15293) MtN21 [Medicago truncatula] NCBI Description 10507 Seq. No. 18956 6.R1040 Contig ID zzp700831260.hl 5'-most EST BLASTX Method g3482913 NCBI GI 172 BLAST score 2.0e-12 E value 108 Match length 33 % identity NCBI Description (AC003970) Similar to MtN21, gi 2598575, Megicago truncatula nodulation induced gene [Arabidopsis thaliana] 10508 Seq. No. 18963 1.R1040 Contig ID LIB3049-002-Q1-E1-C4 5'-most EST BLASTX Method q2827536 NCBI GI 425 BLAST score 2.0e-41 E value 121 Match length 66 % identity (AL021633) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 10509 Contig ID 18964 1.R1040 5'-most EST LIB3106-111-Q1-K1-E2 10510 Seq. No. Contig ID 18964 2.R1040 wrg700789867.h2 5'-most EST Seq. No. 10511 18967 1.R1040 Contig ID

5'-most EST LIB30 $\overline{4}$ 9-002-Q1-E1-C9

Method BLASTX
NCBI GI g3953461
BLAST score 303
E value 2.0e-27



```
Match length
                  101
                  56
% identity
                  (AC002328) F20N2.6 [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  10512
                  18968 1.R1040
Contig ID
                  awf700838162.hl
5'-most EST
                  BLASTX
Method
                  q626464
NCBI GI
                  181
BLAST score
E value
                  5.0e-13
                  68
Match length
% identity
                  VPS27 protein - yeast (Saccharomyces cerevisiae)
NCBI Description
                  >gi_496723_emb_CAA54574_ (X77395) N2038 [Saccharomyces
                  cerevisiae] >gi_1302478_emb CAA96282_ (Z71620) ORF YNR006w
                   [Saccharomyces cerevisiae]
                  10513
Seq. No.
                  18968 2.R1040
Contig ID
                  LIB3065-006-Q1-N1-F3
5'-most EST
Method
                  BLASTX
NCBI GI
                  q626464
BLAST score
                  180
                  2.0e-13
E value
Match length
                  68
% identity
                  VPS27 protein - yeast (Saccharomyces cerevisiae)
NCBI Description
                   >gi_496723_emb_CAA54574_ (X77395) N2038 [Saccharomyces
                   cerevisiae] >gi_1302478_emb_CAA96282_ (Z71620) ORF YNR006w
                   [Saccharomyces cerevisiae]
Seq. No.
                   10514
                   18970 1.R1040
Contig ID
                  LIB3170-023-Q1-K1-H9
5'-most EST
Method
                  BLASTX
NCBI GI
                   g732311
```

BLAST score 158 2.0e-10 E value

100 Match length % identity 38 NCBI Description

HYPOTHETICAL 15.7 KD PROTEIN IN PBPD-COMA INTERGENIC REGION (ORF2) >gi_1075880_pir__B55220 pbpD 3'-region hypothetical protein - Bacillus subtilis >gi_520537 (U11882) unknown [Bacillus subtilis] >gi_1934785_emb_CAB07916_ (Z93933) unknown [Bacillus subtilis] >gi_2635646_emb_CAB15139 (Z99120) alternate gene name: yugD [Bacillus subtilis]

10515 Seq. No.

18979 1.R1040 Contig ID

LIB3170-017-Q1-K1-H9 5'-most EST

BLASTX Method g3831454 NCBI GI 198 BLAST score 4.0e-15 E value 74 Match length % identity 57



```
(AC005700) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  10516
Seq. No.
Contig ID
                  18988 1.R1040
5'-most EST
                  LIB3051-024-Q1-K1-E3
Seq. No.
                  10517
                  18991 1.R1040
Contig ID
                  LIB3170-024-Q1-K1-A3
5'-most EST
                  BLASTX
Method
NCBI GI
                   g1001650
BLAST score
                   183
                   3.0e-13
E value
                   70
Match length
% identity
                  (D64002) hypothetical protein [Synechocystis sp.]
NCBI Description
                   10518
Seq. No.
                   18992 1.R1040
Contig ID
5'-most EST
                   dpv701099639.h1
                   BLASTX
Method
NCBI GI
                   q4539302
                   384
BLAST score
                   4.0e-37
E value
                   94
Match length
                   79
% identity
                  (AL049480) putative protein [Arabidopsis thaliana]
NCBI Description
                   10519
Seq. No.
Contig ID
                   18995 1.R1040
                   LIB3049-001-Q1-E1-H12
5'-most EST
Seq. No.
                   10520
                   19001 1.R1040
Contig ID
                   LIB3049-001-Q1-E1-H7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g4309732
BLAST score
                   229
                   1.0e-18
E value
Match length
                   128
% identity
                   41
                   (AC006439) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10521
Seq. No.
                   19003 1.R1040
Contig ID
                   uC-gmflminsoy059c11b1
5'-most EST
                   BLASTX
Method
                   g3355480
NCBI GI
BLAST score
                   715
                   2.0e-75
E value
Match length
                   277
                   52
% identity
                   (AC004218) Medicago nodulin N21-like protein [Arabidopsis
NCBI Description
                   thaliana]
```

1837

10522

19013 1.R1040

Seq. No.

Contig ID



5'-most EST LIB3106-036-Q1-K1-A11

Method BLASTX
NCBI GI g3643606
BLAST score 298
E value 1.0e-26
Match length 158
% identity 38

NCBI Description (AC005395) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10523

Contig ID 19027 1.R1040

5'-most EST LIB3049-001-Q1-E1-H1

Seq. No. 10524

Contig ID 19030_1.R1040

5'-most EST LIB3049-001-Q1-E1-E10

Seq. No. 10525

Contig ID 19038_1.R1040 5'-most EST gsv701053435.h1

Method BLASTN
NCBI GI g3204107
BLAST score 215
E value 1.0e-117
Match length 475
% identity 86

NCBI Description Cicer arietinum mRNA for putative cytidine or

deoxycytidylate deaminase, partial

Seq. No. 10526

Contig ID 19038_2.R1040 5'-most EST leu701156293.h1

Method BLASTN
NCBI GI g3204107
BLAST score 323
E value 0.0e+00
Match length 539
% identity 90

NCBI Description Cicer arietinum mRNA for putative cytidine or

deoxycytidylate deaminase, partial

Seq. No. 10527

Contig ID 19038 4.R1040

5'-most EST LIB3051-038-Q1-K1-A12

Method BLASTN
NCBI GI g3204107
BLAST score 255
E value 1.0e-141
Match length 427
% identity 89

NCBI Description Cicer arietinum mRNA for putative cytidine or

deoxycytidylate deaminase, partial

Seq. No. 10528

Contig ID 19045 1.R1040

5'-most EST LIB3049-001-Q1-E1-C1



Seq. No. 1052

Contig ID 19048_1.R1040 5'-most EST kl1701208113.h1

Seq. No. 10530

Contig ID 19050 1.R1040

5'-most EST LIB3049-001-Q1-E1-C4

Method BLASTX
NCBI GI g3367537
BLAST score 148
E value 8.0e-10
Match length 70
% identity 49

NCBI Description (AC004392) Contains similarity to ANK repeat region of

Fowlpox virus BamHI-orf7 protein homolog C18F10.7 gi_485107 from Caenorhabditis elegans cosmid gb_U00049. This gene is continued from unannotated gene on BAC F19K23 gb_AC000375.

[Arabid

Seq. No. 10531

Contig ID 19063 1.R1040

5'-most EST LIB3049-001-Q1-E1-D6

Seq. No. 10532

Contig ID 19065_1.R1040
5'-most EST g4306618
Method BLASTX
NCBI GI g1084391
BLAST score 432
E value 3.0e-42

Match length 124 % identity 63

NCBI Description endo-1,4-beta-D-glucanase, xyloglucan-specific (clone NXG2)

- common nasturtium (fragment) >gi_311837_emb_CAA48325_

(X68255) cellulase [Tropaeolum majus]

Seq. No. 10533

Contig ID 19067 1.R1040

5'-most EST LIB3049-001-Q1-E1-A7

Seq. No. 10534

Contig ID 19069 1.R1040

5'-most EST LIB3170-020-Q1-K1-H6

Seq. No. 10535

Contig ID 19071 1.R1040

5'-most EST LIB3049-001-Q1-E1-B12

Seq. No. 10536

Contig ID 19093_1.R1040

5'-most EST LIB3050-022-Q1-K1-A7

Seq. No. 10537

Contig ID 19093_2.R1040

5'-most EST LIB3049-008-Q1-E1-A4

Seq. No. 10538



```
19094 1.R1040
Contig ID
                  LIB3049-009-Q1-E1-C3
5'-most EST
                  10539
Seq. No.
                  19095 1.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy058e07b1
                  BLASTX
Method
NCBI GI
                  q1944407
                  493
BLAST score
                  5.0e-49
E value
Match length
                  312
% identity
                  41
NCBI Description (D86988) KIAA0221 [Homo sapiens]
                  10540
Seq. No.
                  19095 2.R1040
Contig ID
5'-most EST
                  rlr700902272.h1
                  BLASTX
Method
NCBI GI
                  g1944407
BLAST score
                  512
                  5.0e-52
E value
                  124
Match length
% identity
                  75
NCBI Description (D86988) KIAA0221 [Homo sapiens]
                  10541
Seq. No.
Contig ID
                   19096 1.R1040
                  LIB3049-010-Q1-E1-A7
5'-most EST
                   10542
Seq. No.
Contig ID
                   19097 1.R1040
                  LIB3049-010-Q1-E1-H10
5'-most EST
                   10543
Seq. No.
                   19106 1.R1040
Contig ID
5'-most EST
                   g5175461
Method
                   BLASTX
                   g4006875
NCBI GI
BLAST score
                   179
                   1.0e-12
E value
Match length
                   73
% identity
                   59
                  (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
                   10544
Seq. No.
                   19106 2.R1040
Contig ID
                   LIB3049-012-Q1-E1-H4
5'-most EST
                   10545
```

Seq. No.

19106 3.R1040 Contig ID 5'-most EST wvk700682411.h2

Seq. No.

10546

10547

19111 1.R1040 Contig ID

5'-most EST uC-qmrominsoy173e11b1

Seq. No.



Contig ID 19116_1.R1040

5'-most EST LIB3170-026-Q1-K1-D1

Method BLASTX
NCBI GI g2088647
BLAST score 756
E value 3.0e-80
Match length 279
% identity 56

NCBI Description (AF002109) hypothetical protein [Arabidopsis thaliana]

>gi 3158394 (AF036340) LRR-containing F-box protein

[Arabidopsis thaliana]

Seq. No. 10548

Contig ID 19122 1.R1040

5'-most EST LIB3170-026-Q1-K1-G3

Seq. No. 10549

Contig ID 19125_1.R1040

5'-most EST LIB3049-022-Q1-E1-H12

Method BLASTX
NCBI GI g2739369
BLAST score 594
E value 1.0e-61
Match length 133

% identity 81

NCBI Description (AC002505) putative pectinesterase [Arabidopsis thaliana]

Seq. No. 10550

Contig ID 19132_1.R1040

5'-most EST LIB3106-100-Q1-K1-A3

Method BLASTX
NCBI GI g4538929
BLAST score 326
E value 6.0e-30
Match length 113
% identity 60

NCBI Description (AL049483) putative nucleic acid binding protein

[Arabidopsis thaliana]

Seq. No. 10551

Contig ID 19139 1.R1040

5'-most EST LIB3049-029-Q1-E1-C6

Seq. No. 10552

Contig ID 19139 2.R1040

5'-most EST jC-gmle01810093e08a1

Seq. No. 10553

Contig ID 19143 1.R1040

5'-most EST LIB3049-029-Q1-E1-E5

Seq. No. 10554

Contig ID 19144 1.R1040

5'-most EST LIB3049-029-Q1-E1-F11

Method BLASTX
NCBI GI g4191783
BLAST score 188



```
7.0e-14
E value
                   55
Match length
                   60
% identity
                   (AC005917) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10555
                   19144 2.R1040
Contig ID
                   uC-qmflminsoy025g03b1
5'-most EST
                   10556
Seq. No.
                   19145 1.R1040
Contig ID
                   LIB30\overline{4}9-029-Q1-E1-F5
5'-most EST
                   BLASTX
Method
                   q3687460
NCBI GI
                   127
BLAST score
                   6.0e-11
E value
Match length
                   99
% identity
                   (AL031798) putative leucine permease transcriptional
NCBI Description
                   regulator. [Schizosaccharomyces pombe]
                   10557
Seq. No.
                   19146 1.R1040
Contig ID
                   LIB3167-077-P1-K2-A3
5'-most EST
                   BLASTX
Method
                   q3927838
NCBI GI
BLAST score
                   416
                   1.0e-40
E value
                   117
Match length
% identity
                   (AC005727) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10558
                   19146 2.R1040
Contig ID
                   LIB3049-029-Q1-E1-H12
 5'-most EST
                   BLASTX
Method
                   g3927838
NCBI GI
                   165
BLAST score
                   2.0e-11
E value
Match length
                   40
 % identity
                   72
                   (AC005727) unknown protein [Arabidopsis thaliana]
 NCBI Description
                   10559
 Seq. No.
                   19148 1.R1040
 Contig ID
                   LIB3049-030-Q1-E1-F4
 5'-most EST
                    10560
 Seq. No.
                    19149 1.R1040
 Contig ID
                    LIB3109-011-Q1-K1-F2
 5'-most EST
                    BLASTX
 Method
                    g2795806
 NCBI GI
                    287
 BLAST score
                    7.0e-33
 E value
                    152
 Match length
                    55
 % identity
```

1842

NCBI Description (AC003674) unknown protein [Arabidopsis thaliana]



10561 Seq. No. Contig ID 19150 1.R1040 5'-most EST uaw700666775.hl Method BLASTX g3297818 NCBI GI BLAST score 271 6.0e-24 E value 100 Match length 45 % identity NCBI Description (AL031032) putative protein [Arabidopsis thaliana] 10562 Seq. No. 19150 2.R1040 Contig ID zzp700830822.h1 5'-most EST Method BLASTX g3297818 NCBI GI BLAST score 149 3.0e-19 E value 99 Match length % identity NCBI Description (AL031032) putative protein [Arabidopsis thaliana] 10563 Seq. No. 19152 1.R1040 Contig ID LIB3049-033-Q1-E1-C8 5'-most EST 10564 Seq. No. 19153 1.R1040 Contig ID 5'-most EST LIB3049-033-Q1-E1-G12 10565 Seq. No. 19160 1.R1040 Contig ID LIB3170-031-Q1-K1-A10 5'-most EST 10566 Seq. No. Contig ID 19167 1.R1040 LIB3170-085-Q1-J1-A7 5'-most EST Method BLASTX q3980412 NCBI GI BLAST score 481 2.0e-53 E value Match length 205 16 % identity (AC004561) pumilio-like protein [Arabidopsis thaliana] NCBI Description 10567 Seq. No. 19171 1.R1040 Contig ID LIB3093-022-Q1-K1-D9 5'-most EST

Method BLASTX
NCBI GI g3924612
BLAST score 540
E value 5.0e-55
Match length 198
% identity 61

NCBI Description (AF069442) mitochondrial elongation factor Tu [Arabidopsis thaliana] >gi_4263511_gb_AAD15337_ (AC004044) mitochondrial



elongation factor Tu [Arabidopsis thaliana]

10568 Seq. No. Contig ID 19175 1.R1040 LIB3049-047-Q1-E1-G8 5'-most EST Method BLASTX q1076715 NCBI GI 278 BLAST score E value 2.0e-24 92 Match length 53 % identity abscisic acid-induced protein HVA22 - barley >gi_404589 NCBI Description (L19119) A22 [Hordeum vulgare] Seq. No. 10569 Contig ID 19180 1.R1040 uC-gmrominsoy186e08b1 5'-most EST Method BLASTX q3142301 NCBI GI 295 BLAST score 2.0e-26 E value 101 Match length % identity (AC002411) Contains similarity to neural cell adhesion NCBI Description molecule 2, large isoform precursor gb M76710 from Xenopus laevis, and beta transducin from S. cerevisiae gb_Q05946. ESTs gb_N65081 gb_Z30910, gb_Z34190, gb_Z34611, gb_R30101, gb H3630 10570 Seq. No. 19182 1.R1040 Contig ID LIB3049-051-Q1-E1-F5 5'-most EST Method BLASTX g2262159 NCBI GI 961 BLAST score 1.0e-104 E value 265 Match length 66 % identity (AC002329) predicted protein similar to S.pombe protein NCBI Description C5H10.03 [Arabidopsis thaliana] 10571 Seq. No. 19185 1.R1040 Contig ID LIB3170-035-Q1-K1-A12 5'-most EST BLASTX Method g4539351 NCBI GI 289 BLAST score 9.0e-26 E value 160 Match length 39 % identity (AL035539) putative protein [Arabidopsis thaliana] NCBI Description 10572 Seq. No.

5'-most EST LIB3106-100-Q1-K1-H11 Method BLASTX

19201_1.R1040

Method BLASTX NCBI GI g4105772

Contig ID



```
740
BLAST score
                  2.0e-78
E value
                  180
Match length
                  70
% identity
NCBI Description (AF049917) PGP9B [Petunia x hybrida]
Seq. No.
                  10573
                  19203 1.R1040
Contig ID
                  kl1701214886.h1
5'-most EST
                  10574
Seq. No.
                  19212 1.R1040
Contig ID
                  g4396976
5'-most EST
                  BLASTX
Method
                  g3334230
NCBI GI
                  1143
BLAST score
                  1.0e-125
E value
Match length
                  441
                  50
% identity
NCBI Description D-HYDANTOINASE (DIHYDROPYRIMIDINASE) (DHPASE) >gi 2828803
                   (U84197) D-hydantoinase [Pseudomonas putida]
                  10575
Seq. No.
                  19212 2.R1040
Contig ID
                  LIB3109-038-Q1-K1-F5
5'-most EST
                  BLASTX
Method
                  g3334230
NCBI GI
BLAST score
                   450
                   9.0e-80
E value
                   315
Match length
                   51
% identity
NCBI Description D-HYDANTOINASE (DIHYDROPYRIMIDINASE) (DHPASE) >gi_2828803
                   (U84197) D-hydantoinase [Pseudomonas putida]
                   10576
Seq. No.
Contig ID
                   19212 3.R1040
                   uC-gmflminsoy118d03b1
5'-most EST
                   10577
Seq. No.
                   19215 1.R1040
Contig ID
5'-most EST
                   LIB3050-028-Q1-E1-G8
Seq. No.
                   10578
Contig ID
                   19216 1.R1040
                   LIB3139-044-P1-N1-F12
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2146775
BLAST score
                   438
                   6.0e-86
E value
Match length
                   342
% identity
                   52
                   serine acetyltransferase (EC 2.3.1.30) SAT1 - Arabidopsis
NCBI Description
                   thaliana >gi_926939 (L42212) serine acetyltransferase
```

Seq. No. 10579

Contig ID 19218 1.R1040

1845

[Arabidopsis thaliana]



```
LIB3050-028-Q1-E1-H10
5'-most EST
                  BLASTX
Method
                  g4314361
NCBI GI
                  216
BLAST score
                  2.0e-17
E value
Match length
                  44
% identity
                  (AC006340) unknown protein [Arabidopsis thaliana]
NCBI Description
                  10580
Seq. No.
                   19218 2.R1040
Contig ID
                   uC-gmropic100h07b1
5'-most EST
                   BLASTX
Method
                   g4314361
NCBI GI
                   222
BLAST score
                   5.0e-18
E value
Match length
                   45
                   96
% identity
NCBI Description (AC006340) unknown protein [Arabidopsis thaliana]
                   10581
Seq. No.
                   19223 1.R1040
Contig ID
                   jC-gmle01810053c09a1
5'-most EST
                   BLASTX
Method
                   g4006899
NCBI GI
                   482
BLAST score
                   2.0e-48
E value
Match length
                   138
                   64
% identity
NCBI Description (Z99708) putative protein [Arabidopsis thaliana]
                   10582
Seq. No.
                   19223 3.R1040
Contig ID
                   rca701001718.hl
5'-most EST
                   10583
Seq. No.
                   19229 2.R1040
Contig ID
5'-most EST
                   ssr700556492.hl
                   10584
Seq. No.
                   19229 3.R1040
Contig ID
                   6HC - 02 - Q1 - B1 - E8
5'-most EST
Seq. No.
                   10585
                   19229 4.R1040
Contig ID
                   LIB3139-022-P1-N1-D5
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1665817
BLAST score
                   175
                   2.0e-12
E value
Match length
                   68
                   47
 % identity
                   (D87466) Similar to S.cerevisiae hypothetical protein L3111
NCBI Description
                    (S59316) [Homo sapiens]
```

10586

19231 1.R1040

Seq. No.

Contig ID



```
fC-gmst700652392f1
5'-most EST
Seq. No.
                  10587
Contig ID
                  19234 1.R1040
                  asn701135274.h1
5'-most EST
                  BLASTX
Method
                  g2464852
NCBI GI
BLAST score
                  316
                  4.0e-29
E value
Match length
                  124
% identity
                  (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
                  10588
Seq. No.
Contig ID
                  19235 1.R1040
5'-most EST
                  LIB3139-041-P1-N1-F8
Method
                  BLASTX
NCBI GI
                   q3157951
BLAST score
                   710
                   3.0e-93
E value
                   217
Match length
% identity
                   85
                   (AC002131) Contains similarity to vesicle trafficking
NCBI Description
                   protein gb U91538 from Mus musculus. ESTs gb F15494 and
                   gb F14097 come from this gene. [Arabidopsis thaliana]
                   10589
Seq. No.
Contig ID
                   19235 2.R1040
5'-most EST
                   hrw701062691.h1
                   BLASTX
Method
                   g3157951
NCBI GI
BLAST score
                   284
                   2.0e-25
E value
                   64
Match length
                   84
% identity
                   (AC002131) Contains similarity to vesicle trafficking
NCBI Description
                   protein gb U91538 from Mus musculus. ESTs gb F15494 and
                   gb F14097 come from this gene. [Arabidopsis thaliana]
                   10590
Seq. No.
                   19235 3.R1040
Contig ID
                   zsg701117409.h1
5'-most EST
                   10591
Seq. No.
                   19237 1.R1040
Contig ID
5'-most EST
                   uC-gmflminsoy022d12b1
                   BLASTX
Method
                   g4218062
NCBI GI
BLAST score
                   374
E value
                   1.0e-35
Match length
                   172
% identity
                   42
```

NCBI Description

Seq. No.

10592

Contig ID 19240 1.R1040

5'-most EST LIB3072-051-Q1-E1-A9

(AB015906) actin-related protein [Homo sapiens]



```
BLASTX
Method
                  q4539348
NCBI GI
                  934
BLAST score
                  1.0e-101
E value
                  247
Match length
                  68
% identity
NCBI Description (AL035539) putative pollen allergen [Arabidopsis thaliana]
                  10593
Seq. No.
                  19240 2.R1040
Contig ID
                  LIB3139-085-P1-N1-D11
5'-most EST
                  BLASTX
Method
                  g4539348
NCBI GI
                  198
BLAST score
                   8.0e-26
E value
                   99
Match length
% identity
                  (AL035539) putative pollen allergen [Arabidopsis thaliana]
NCBI Description
                   10594
Seq. No.
                   19240 3.R1040
Contig ID
                   uC-gmrominsoy169c10b1
5'-most EST
                   BLASTX
Method
                   g4539348
NCBI GI
                   347
BLAST score
                   1.0e-32
E value
                   86
Match length
                   69
% identity
                  (AL035539) putative pollen allergen [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10595
                   19242 1.R1040
Contig ID
5'-most EST
                   LIB3050-002-Q1-E1-D5
Seq. No.
                   10596
                   19242 2.R1040
Contig ID
                   LIB3107-061-Q1-K1-F12
5'-most EST
                   10597
Seq. No.
                   19246 1.R1040
Contig ID
                   jC-gmst02400036g10d2
5'-most EST
Seq. No.
                   10598
                   19247 1.R1040
Contig ID
                   jC-gmf102220061a10a1
 5'-most EST
Method
                   BLASTX
                   q4006875
NCBI GI
BLAST score
                   323
                   8.0e-30
E value
                   113
Match length
 % identity
                   (Z99707) putative protein [Arabidopsis thaliana]
NCBI Description
                   10599
 Seq. No.
```

19248 1.R1040

dpv701102923.h1

Contig ID

5'-most EST



```
10600
Seq. No.
                  19248 2.R1040
Contig ID
5'-most EST
                  LIB3107-036-Q1-K1-D3
                  BLASTX
Method
NCBI GI
                  g2245060
BLAST score
                  280
                  9.0e-25
E value
                  76
Match length
% identity
NCBI Description (Z97342) allergen homolog [Arabidopsis thaliana]
                  10601
Seq. No.
                  19250 1.R1040
Contig ID
                  LIB3050-028-Q1-E1-C4
5'-most EST
                   10602
Seq. No.
                  19252 1.R1040
Contig ID
5'-most EST
                  wvk700685013.hl
Method
                  BLASTX
                   g3281850
NCBI GI
                   788
BLAST score
                   3.0e-84
E value
                   175
Match length
                   84
% identity
NCBI Description (AL031004) monogalactosyldiacylglycerol synthase - like
                   protein [Arabidopsis thaliana]
Seq. No.
                   10603
                   19255 1.R1040
Contig ID
5'-most EST
                   uC-gmflminsoy080g12b1
Method
                   BLASTX
                   q2213610
NCBI GI
BLAST score
                   245
E value
                   3.0e-20
Match length
                   150
                   39
% identity
                  (AC000103) F21J9.4 [Arabidopsis thaliana]
NCBI Description
                   10604
Seq. No.
Contig ID
                   19257 1.R1040
                   LIB3050-027-Q1-E1-H4
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4314390
BLAST score
                   396
                   4.0e-38
E value
Match length
                   142
% identity
                   54
                   (AC006232) putative calcium binding protein [Arabidopsis
NCBI Description
                   thaliana]
                   10605
Seq. No.
                   19257 2.R1040
Contig ID
5'-most EST
                   zsg701124230.h1
                   BLASTX
Method
```

g4314390

2.0e-31

337

NCBI GI

E value

BLAST score



Match length % identity

(AC006232) putative calcium binding protein [Arabidopsis NCBI Description

thaliana]

Seq. No. 10606

19257 3.R1040 Contig ID uaw700664521.hl 5'-most EST

BLASTX Method q4314390 NCBI GI BLAST score 200 2.0e-15 E value 80 Match length % identity

(AC006232) putative calcium binding protein [Arabidopsis NCBI Description

thaliana]

10607 Seq. No.

19257 4.R1040 Contig ID

uC-gmflminsoy056h10b1 5'-most EST

10608 Seq. No.

19259 1.R1040 Contig ID

LIB3050-027-Q1-E1-H6 5'-most EST

10609 Seq. No.

19260 1.R1040 Contig ID

uC-gmrominsoy299e06b1 5'-most EST

BLASTX Method q1743354 NCBI GI 1475 BLAST score 1.0e-164 E value Match length 484 57

% identity

(Y09876) aldehyde dehydrogenase (NAD+) [Nicotiana tabacum] NCBI Description

10610 Seq. No.

19262 1.R1040 Contig ID

uC-gmrominsoy223a06b1 5'-most EST

BLASTX Method g4063742 NCBI GI BLAST score 218 E value 4.0e-17 Match length 153 % identity 38

(AC005851) putative phaseolin G-box binding protein NCBI Description

[Arabidopsis thaliana]

10611 Seq. No.

19262 2.R1040 Contig ID rlr700900662.h1 5'-most EST

Seq. No.

19263 1.R1040 Contig ID

 $LIB30\overline{5}1-067-Q1-K1-G10$ 5'-most EST

10612

10613 Seq. No.



Contig ID 19264 1.R1040 5'-most EST dpv701102455.h1

Seq. No. 10614

Contig ID 19266_1.R1040

5'-most EST LIB3139-017-P1-N1-C1

Seq. No. 10615

Contig ID 19267_1.R1040 5'-most EST zhf700952901.h1

Method BLASTX
NCBI GI g2654868
BLAST score 306
E value 4.0e-52
Match length 174
% identity 66

NCBI Description (AF015301) RbohAp108 [Arabidopsis thaliana]

Seq. No. 10616

Contig ID 19268 1.R1040

5'-most EST jC-gmf102220097f01a1

Seq. No. 10617

Contig ID 19268 2.R1040

5'-most EST LIB3051-027-Q1-K1-G5

Seq. No. 10618

Contig ID 19274_1.R1040

5'-most EST LIB3050-027-Q1-E1-F3

Method BLASTN
NCBI GI g287563
BLAST score 77
E value 6.0e-35
Match length 327
% identity 91

NCBI Description Vigna radiata arg2 mRNA

Seq. No. 10619

Contig ID 19274_2.R1040

5'-most EST LIB3050-006-Q1-E1-G12

Method BLASTN
NCBI GI g287563
BLAST score 90
E value 1.0e-42
Match length 338
% identity 90

NCBI Description Vigna radiata arg2 mRNA

Seq. No. 10620

Contig ID 19275_1.R1040

5'-most EST LIB3139-084-P1-N1-H2

Seq. No. 10621

Contig ID 19275_2.R1040 5'-most EST asn701132347.h1

Seq. No. 10622



```
19278 1.R1040
Contig ID
                   kmv70\overline{0}737840.h1
5'-most EST
Seq. No.
                   10623
Contig ID
                   19288 1.R1040
                   sat701004275.hl
5'-most EST
Method
                   BLASTX
                   g2435511
NCBI GI
BLAST score
                   449
E value
                   3.0e-44
Match length
                   144
                   58
% identity
                  (AF024504) contains similarity to prolyl 4-hydroxylase
NCBI Description
                   alpha subunit [Arabidopsis thaliana]
Seq. No.
                   10624
                   19288 2.R1040
Contig ID
                   jex700906632.hl
5'-most EST
Seq. No.
                   10625
Contig ID
                   19293 1.R1040
                   hrw701058172.h1
5'-most EST
Method
                   BLASTX
                   g4455351
NCBI GI
BLAST score
                   653
E value
                   5.0e-68
Match length
                   183
% identity
                   68
                   (AL035524) putative protein [Arabidopsis thaliana]
NCBI Description
                   10626
Seq. No.
                   19293 2.R1040
Contig ID
                   LIB3051-037-Q1-K1-A1
5'-most EST
Method
                   BLASTX
                   g4455351
NCBI GI
BLAST score
                   298
                   8.0e-27
E value
Match length
                   82
% identity
                   (AL035524) putative protein [Arabidopsis thaliana]
NCBI Description
                   10627
Seq. No.
                   19299 1.R1040
Contig ID
                   leu701148775.h1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g1651451
                   517
```

Method BLASTX
NCBI GI g165145
BLAST score 517
E value 2.0e-52
Match length 137
% identity 70

NCBI Description (D90730) Asparaginyl-tRNA synthetase (EC 6.1.1.22) (asparagine-tRNA ligase) (asnRS). [Escherichia coli]

Seq. No. 10628

Contig ID 19305 1.R1040

5'-most EST LIB3050-021-Q1-K1-E2

Method BLASTX

BLAST score

Match length

E value

235 1.0e-19

95



```
q1762933
NCBI GI
BLAST score
                  476
                  2.0e-47
E value
                  203
Match length
% identity
NCBI Description (U66263) tumor-related protein [Nicotiana tabacum]
                  10629
Seq. No.
                  19305 2.R1040
Contig ID
                  leu701155692.h1
5'-most EST
Method
                  BLASTX
                  q1762933
NCBI GI
                  291
BLAST score
                  3.0e - 26
E value
                  95
Match length
                   63
% identity
NCBI Description (U66263) tumor-related protein [Nicotiana tabacum]
Seq. No.
                  10630
                  19306 1.R1040
Contig ID
                  LIB3050-027-Q1-E1-C8
5'-most EST
Seq. No.
                   10631
                   19308 1.R1040
Contig ID
                  LIB3109-031-Q1-K1-B3
5'-most EST
Seq. No.
                   10632
                   19311 1.R1040
Contig ID
                   LIB3170-065-Q1-K1-C8
5'-most EST
                   BLASTX
Method
                   q3367522
NCBI GI
                   478
BLAST score
E value
                   5.0e-48
                   126
Match length
% identity
                   (AC004392) EST gb_T04691 comes from this gene. [Arabidopsis
NCBI Description
                   thaliana]
                   10633
Seq. No.
                   19313 1.R1040
Contig ID
5'-most EST
                   gsv701052993.h1
Method
                   BLASTX
                   q1688233
NCBI GI
                   366
BLAST score
                   2.0e-34
E value
Match length
                   176
% identity
                   (U77655) DNA binding protein homolog [Solanum tuberosum]
NCBI Description
                   10634
Seq. No.
                   19313 2.R1040
Contig ID
                   LIB3050-027-Q1-E1-B7
 5'-most EST
Method
                   BLASTX
                   q1688233
NCBI GI
```



```
% identity
                  (U77655) DNA binding protein homolog [Solanum tuberosum]
NCBI Description
                  10635
Seq. No.
                  19314 1.R1040
Contig ID
                  LIB3050-027-Q1-E1-B8
5'-most EST
Seq. No.
                  10636
                  19321 1.R1040
Contig ID
5'-most EST
                  LIB3051-104-Q1-K1-A3
Method
                  BLASTX
                  g125722
NCBI GI
BLAST score
                  502
                  1.0e-50
E value
Match length
                  194
% identity
                  58
                  KUNITZ-TYPE TRYPSIN INHIBITOR KTI1 PRECURSOR
NCBI Description
                  >qi 81814 pir JQ1091 trypsin inhibitor KTi1 (Kunitz) -
                   soybean >qi 256635 bbs 115028 (S45035) Kunitz trypsin
                   inhibitor KTil [soybeans, Peptide, 203 aa] [Glycine max]
٠<u>٠</u>,
Seq. No.
                   10637
Contig ID
                   19322 1.R1040
                   uC-gmflminsoy049c02b1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3033386
BLAST score
                   267
E value
                   1.0e-22
Match length
                   223
% identity
                   32
                  (AC004238) RING3-like protein [Arabidopsis thaliana]
NCBI Description
                   10638
Seq. No.
Contig ID
                   19322 2.R1040
5'-most EST
                   uC-gmrominsoy260e09b1
                   10639
Seq. No.
                   19322 5.R1040
Contig ID
                   LIB3139-011-P1-N1-G11
5'-most EST
Seq. No.
                   10640
                   19326 1.R1040
Contig ID
                   txt700735881.h1
5'-most EST
                   10641
Seq. No.
Contig ID
                   19329_1.R1040
                   g4289596
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g1050956
BLAST score 314
E value 2.0e-28
Match length 242
% identity 36

NCBI Description (U26160) 43 kDa secreted glycoprotein precursor

[Paracoccidioides brasiliensis] >gi_1588394_prf__2208385A

glycoprotein gp43 [Paracoccidioides brasiliensis]



```
10642
Seq. No.
                  19331 1.R1040
Contig ID
5'-most EST
                  uC-gmronoir065g05b1
Method
                  BLASTN
NCBI GI
                  g1684852
                  392
BLAST score
                  0.0e+00
E value
                  630
Match length
                  91
% identity
                  Phaseolus vulgaris PvSR7 protein mRNA, partial cds
NCBI Description
                  10643
Seq. No.
```

19331 2.R1040 Contig ID uC-gmrominsoy217b05b1 5'-most EST Method BLASTX

NCBI GI g4033469 BLAST score 174 E value 2.0e-12 Match length 39 % identity 74

ARGININE/SERINE-RICH SPLICING FACTOR RSP41 NCBI Description

>qi 1707370 emb CAA67799 (X99436) splicing factor

[Arabidopsis thaliana]

Seq. No. 10644

Contig ID 19332 1.R1040

LIB3050-026-Q1-K1-H7 5'-most EST

Seq. No. 10645

19333 1.R1040 Contig ID $jex70\overline{0}905478.h1$ 5'-most EST

10646 Seq. No.

19334 1.R1040 Contig ID

5'-most EST LIB3050-026-Q1-K1-H9

10647 Seq. No.

19335 1.R1040 Contig ID 5'-most EST uC-gmropic093d12b1

BLASTX Method g3023961 NCBI GI 967 BLAST score 1.0e-105 E value 212 Match length 84 % identity

HOMEOBOX PROTEIN KNOTTED-1-LIKE >gi 1946222 emb CAA96512 NCBI Description (Z71980) knotted1-like homeobox protein [Malus domestica]

10648 Seq. No.

19335 2.R1040 Contig ID

 $jC-gm\overline{l}e01810025e10d1$ 5'-most EST

BLASTN Method q1946221 NCBI GI 63 BLAST score 9.0e-27 E value Match length 95 92 % identity



NCBI Description M.domestica mRNA for knotted1-like homeobox protein

Seq. No. 10649

Contig ID 19336_1.R1040 5'-most EST dpv701100409.h1

Method BLASTX
NCBI GI g3367578
BLAST score 260
E value 3.0e-22
Match length 107
% identity 45

NCBI Description (AL031135) protein kinase - like protein [Arabidopsis

thaliana]

Seq. No. 10650

Contig ID 19339_1.R1040 5'-most EST uC-gmronoir013g04b1

Seq. No. 10651

Contig ID 19339 2.R1040 5'-most EST ssr700557305.h1

Seq. No. 10652

Contig ID 19342_1.R1040 5'-most EST uC-gmropic029d04b1

Method BLASTX
NCBI GI g1129145
BLAST score 1176
E value 1.0e-129
Match length 362
% identity 66

NCBI Description (X75329) acetyl-CoA C-acyltransferase [Mangifera indica]

Seq. No. 10653

Contig ID 19351 1.R1040

5'-most EST LIB3170-057-Q1-K1-D3

Seq. No. 10654

Contig ID 19351_2.R1040

5'-most EST LIB3107-082-Q1-K1-H5

Method BLASTX
NCBI GI 94512655
BLAST score 852
E value 2.0e-91
Match length 235
% identity 68

NCBI Description (AC007048) putative protein phosphatase 2C [Arabidopsis

thaliana]

Seq. No. 10655

Contig ID 19351 3.R1040

5'-most EST LIB3106-114-Q1-K1-D12

Method BLASTX
NCBI GI g2842482
BLAST score 274
E value 5.0e-24
Match length 85



```
% identity
                   (AL021749) protein phosphatase 2C-like protein [Arabidopsis
NCBI Description
                  thaliana]
                  10656
Seq. No.
                  19351 4.R1040
Contig ID
                  LIB3139-073-P1-N1-H6
5'-most EST
                  BLASTX
Method
                  g2842482
NCBI GI
BLAST score
                   363
                   2.0e-34
E value
Match length
                   95
                   74
% identity
                  (ALO21749) protein phosphatase 2C-like protein [Arabidopsis
NCBI Description
                   thaliana]
                   10657
Seq. No.
                   19351 5.R1040
Contig ID
                   uC-qmronoir068b05b1
5'-most EST
                   BLASTX
Method
                   g4336436
NCBI GI
BLAST score
                   312
                   2.0e-28
E value
                   92
Match length
                   64
% identity
                   (AF092432) protein phosphatase type 2C [Lotus japonicus]
NCBI Description
Seq. No.
                   10658
                   19352 1.R1040
Contig ID
                   kmv70\overline{0}739705.h1
5'-most EST
                   BLASTX
Method
                   g2244866
NCBI GI
                   805
BLAST score
                   5.0e-86
E value
                   223
Match length
                   68
% identity
                  (Z97337) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10659
Seq. No.
                   19353 1.R1040
Contig ID
                   fua701043404.hl
5'-most EST
                   10660
Seq. No.
                   19354 1.R1040
Contig ID
5'-most EST
                   LIB3050-026-Q1-K1-G5
Method
                   BLASTX
                   q3738332
NCBI GI
                   195
BLAST score
                   5.0e-17
E value
Match length
                   81
```

% identity 67

(AC005170) putative eukaryotic initiation factor NCBI Description

[Arabidopsis thaliana]

Seq. No.

10661 19354 2.R1040 Contig ID $uxk70\overline{0}672920.h1$ 5'-most EST



Method BLASTX
NCBI GI g3738332
BLAST score 164
E value 1.0e-11
Match length 88
% identity 51

NCBI Description (AC005170) putative eukaryotic initiation factor

[Arabidopsis thaliana]

Seq. No. 10662

Contig ID 19359_1.R1040

5'-most EST LIB3050-026-Q1-K1-H1

Method BLASTX
NCBI GI g3928862
BLAST score 144
E value 5.0e-09
Match length 125
% identity 30

NCBI Description (AF089710) disease resistance protein RPP8 [Arabidopsis

thaliana]

Seq. No. 10663

Contig ID 19360 1.R1040

5'-most EST LIB3051-115-Q1-K1-C4

Method BLASTX
NCBI GI g282963
BLAST score 579
E value 9.0e-60
Match length 133
% identity 81

NCBI Description transforming protein (myb) homolog (clone myb.Ph2) - garden

petunia >gi 20561 emb CAA78387 (Z13997) protein 2 [Petunia

x hybrida]

Seq. No. 10664

Contig ID 19362 1.R1040 5'-most EST uxk700672889.h1

Seq. No. 10665

Contig ID 19364 1.R1040

5'-most EST uC-gmronoir043h01b1

Method BLASTX
NCBI GI g2829862
BLAST score 773
E value 1.0e-129
Match length 443
% identity 52

NCBI Description (AC002396) Similar to glucosyltransferases [Arabidopsis

thaliana]

Seq. No. 10666

Contig ID 19367 1.R1040

5'-most EST LIB3050-026-Q1-K1-E7

Seq. No. 10667

Contig ID 19367 2.R1040 5'-most EST sat701009408.h1

Match length

% identity

177



```
10668
Seq. No.
Contig ID
                  19372 1.R1040
                  LIB3106-064-Q1-K1-G8
5'-most EST
Method
                  BLASTX
                  g4309728
NCBI GI
                  836
BLAST score
                  9.0e-90
E value
                  185
Match length
% identity
                  88
                  (ACOO6439) putative ADP-ribosylation factor [Arabidopsis
NCBI Description
                   thaliana]
                   10669
Seq. No.
Contig ID
                   19373 1.R1040
5'-most EST
                   uC-gmronoir045f02b1
Method
                   BLASTX
NCBI GI
                   g3935183
BLAST score
                   358
E value
                   8.0e-34
Match length
                   222
% identity
                   38
NCBI Description (AC004557) F17L21.26 [Arabidopsis thaliana]
Seq. No.
                   10670
Contig ID
                   19376 1.R1040
                   asn701136271.h1
5'-most EST
Method
                   BLASTX
                   q1843462
NCBI GI
BLAST score
                   403
                   4.0e-39
E value
Match length
                   131
                   56
% identity
                   (L10211) isoliquiritigenin 2'-O-methyltransferase [Medicago
NCBI Description
                   sativa]
                   10671
Seq. No.
Contig ID
                   19376 2.R1040
5'-most EST
                   jC-gmro02910025a04d1
                   BLASTX
Method
NCBI GI
                   g1843462
BLAST score
                   296
                   2.0e-26
E value
                   98
Match length
                   55
% identity
                   (L10211) isoliquiritigenin 2'-O-methyltransferase [Medicago
NCBI Description
                   sativa]
                   10672
Seq. No.
                   19377 1.R1040
Contig ID
5'-most EST
                   asn701133492.hl
                   BLASTX
Method
NCBI GI
                   g4558568
BLAST score
                   467
                   1.0e-46
E value
```



NCBI Description (AC007138) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10673

Contig ID 19384_1.R1040 5'-most EST sat701006111.h1

Seq. No. 10674

Contig ID 19385 1.R1040

5'-most EST LIB3050-026-Q1-K1-D12

Seq. No. 10675

Contig ID 19391 1.R1040

5'-most EST LIB3139-006-P1-N1-D8

Seq. No. 10676

Contig ID 19393 1.R1040 5'-most EST fC-gmle700686111f1

Method BLASTX
NCBI GI g1723242
BLAST score 212
E value 2.0e-16
Match length 173
% identity 35

NCBI Description HYPOTHETICAL 36.8 KD PROTEIN C26A3.16 IN CHROMOSOME I >qi 1177363 emb CAA93239 (Z69240) yeast dsk2 homolog,

ubiquitin-like protein [Schizosaccharomyces pombe]

Seq. No. 10677

Contig ID 19393 2.R1040 5'-most EST fC-qmle700686111b1

Method BLASTX
NCBI GI g1723242
BLAST score 146
E value 4.0e-09
Match length 58
% identity 50

NCBI Description HYPOTHETICAL 36.8 KD PROTEIN C26A3.16 IN CHROMOSOME I >gi 1177363 emb CAA93239 (Z69240) yeast dsk2 homolog,

ubiquitin-like protein [Schizosaccharomyces pombe]

Seq. No. 10678

Contig ID 19400 1.R1040 5'-most EST 6HC-01-Q1-B1-A9

Seq. No. 10679

Contig ID 19402_1.R1040

5'-most EST LIB3050-026-Q1-K1-B12

Method BLASTX
NCBI GI g1706082
BLAST score 347
E value 1.0e-32
Match length 139
% identity 55

NCBI Description SERINE CARBOXYPEPTIDASE II-3 PRECURSOR (CP-MII.3)

>gi_629787_pir__S44191 serine-type carboxypeptidase (EC

3.4.16.1) II-3 - barley >gi_619350_bbs_153536

CP-MII.3=serine carboxypeptidase [Hordeum vulgare=barley,



cv. Alexis, aleurone, Peptide, 516 aa]
>gi_474392_emb_CAA55478_ (X78877) serine carboxylase II-3
[Hordeum vulgare]

Seq. No. 10680

Contig ID 19405_1.R1040 5'-most EST bth700847244.h1

Method BLASTX
NCBI GI g1652105
BLAST score 202
E value 2.0e-15
Match length 84
% identity 51

NCBI Description (D90902) hypothetical protein [Synechocystis sp.]

Seq. No. 10681

Contig ID 19405 2.R1040

5'-most EST LIB3050-025-Q1-K1-B11

Method BLASTX
NCBI GI g1652105
BLAST score 192
E value 2.0e-14
Match length 82
% identity 49

NCBI Description (D90902) hypothetical protein [Synechocystis sp.]

Seq. No. 10682

Contig ID 19408 1.R1040

5'-most EST LIB3170-067-Q1-K1-H2

Method BLASTX
NCBI GI g2245007
BLAST score 256
E value 1.0e-21
Match length 96
% identity 58

NCBI Description (Z97341) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10683

Contig ID 19414_1.R1040

5'-most EST LIB3050-025-Q1-K1-G8

Method BLASTX
NCBI GI g1723878
BLAST score 433
E value 7.0e-43
Match length 143
% identity 50

NCBI Description HYPOTHETICAL 65.3 KD PROTEIN IN MAD1-SCY1 INTERGENIC REGION

>gi_2132541_pir__S64091 probable membrane protein YGL084c yeast (Saccharomyces cerevisiae) >gi_1322607_emb_CAA96789_

(Z72606) ORF YGL084c [Saccharomyces cerevisiae]

Seq. No. 10684

Contig ID 19415 1.R1040

5'-most EST LIB3050-025-Q1-K1-G9

Seq. No. 10685

Contig ID 19417_1.R1040



```
5'-most EST
                  LIB3050-025-Q1-K1-H10
Seq. No.
                  10686
Contig ID
                  19422 1.R1040
                  sat701010659.h1
5'-most EST
                  BLASTX
Method
                  g2244833
NCBI GI
                  255
BLAST score
                  2.0e-21
E value
                  328
Match length
                  29
% identity
NCBI Description (Z97337) centromere protein homolog [Arabidopsis thaliana]
                  10687
Seq. No.
                  19423 1.R1040
Contig ID
                  LIB3109-057-Q1-K1-D5
5'-most EST
                  10688
Seq. No.
                  19424 1.R1040
Contig ID
                  k1170\overline{1}207370.h1
5'-most EST
                  10689
Seq. No.
                  19425 1.R1040
Contig ID
                  dpv701097982.h1
5'-most EST
                  BLASTX
Method
                  g2245093
NCBI GI
BLAST score
                  845
                  1.0e-90
E value
                  249
Match length
                   67
% identity
NCBI Description (Z97343) membrane channel protein [Arabidopsis thaliana]
                   10690
Seq. No.
                   19425 2.R1040
Contig ID
                   jC-gmro02910035b05d1
5'-most EST
Method
                   BLASTN
                   q257237
NCBI GI
BLAST score
                   69
                   2.0e-30
E value
Match length
                   201
% identity
                   84
                  TobRB7=root-specific gene regulator [Nicotiana
NCBI Description
                   tabacum=tobacco, cv Wisconsin 38, Genomic, 3426 nt]
                   >gi_1250027_gb_I15119_I15119 Sequence 1 from patent US
                   5459252 >gi_3966754_gb_AR007270_AR007270 Sequence 1 from
                   patent US
                   10691
Seq. No.
                   19425 3.R1040
Contig ID
                   fua701043434.h1
5'-most EST
Method
                   BLASTX
                   q126959
NCBI GI
BLAST score
                   348
                   6.0e-33
E value
                   105
Match length
% identity
NCBI Description TONOPLAST INTRINSIC PROTEIN, ROOT-SPECIFIC RB7-5A (RT-TIP)
```



>gi_82192_pir__JQ1011 TobRB7-5A protein - common tobacco
>gi_100371_pir__S13719 probable membrane channel protein common tobacco >gi_20011_emb_CAA38634_ (X54855) possible
membrane channel protein [Nicotiana tabacum]

Seq. No. 10692

Contig ID 19431_1.R1040 5'-most EST dpv701102817.h1

Method BLASTX
NCBI GI g3157943
BLAST score 295
E value 1.0e-26
Match length 146
% identity 44

NCBI Description (AC002131) Contains similarity to BAP31 protein gb_X81816

from Mus musculus. [Arabidopsis thaliana]

Seq. No. 10693

Contig ID 19434_1.R1040 5'-most EST rlr700900866.h1

Seq. No. 10694

Contig ID 19434 2.R1040

5'-most EST uC-gmrominsoy230c07b1

Seq. No. 10695

Contig ID 19440_1.R1040 5'-most EST awf700837923.h1

Method BLASTN
NCBI GI g3021356
BLAST score 108
E value 1.0e-53
Match length 164
% identity 91

NCBI Description Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase,

clone GEPI48

Seq. No. 10696

Contig ID 19440_2.R1040

5'-most EST LIB3050-025-Q1-K1-G10

Method BLASTN
NCBI GI g3021356
BLAST score 76
E value 2.0e-34
Match length 116
% identity 91

NCBI Description Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase,

clone GEPI48

Seq. No. 10697

Contig ID 19445 1.R1040

5'-most EST uC-gmrominsoy261h11b1

Seq. No. 10698

Contig ID 19445_2.R1040 5'-most EST hrw701060687.h1



Seq. No. 10699

Contig ID 19447_1.R1040 5'-most EST fua701037771.h1

Method BLASTX
NCBI GI g477819
BLAST score 1102
E value 1.0e-121
Match length 268
% identity 78

NCBI Description mitochondrial processing peptidase (EC 3.4.99.41) beta

chain precursor - potato >gi_410634_bbs_136741 cytochrome c reductase-processing peptidase subunit II, MPP subunit II, P53 [potatoes, var. Marfona, tuber, Peptide Mitochondrial,

530 aa]

Seq. No. 10700

Contig ID 19447 2.R1040

5'-most EST LIB3170-046-Q1-J1-D12

Method BLASTX
NCBI GI g477280
BLAST score 242
E value 2.0e-20
Match length 57
% identity 75

NCBI Description mitochondrial processing peptidase (EC 3.4.99.41) 55K

protein precursor - potato >gi_410633_bbs_136740 cytochrome c reductase-processing peptidase subunit I, MPP subunit I, P55 [potatoes, var. Marfona, tuber, Peptide Mitochondrial,

534 aa]

Seq. No. 10701

Contig ID 19452_1.R1040

5'-most EST LIB3050-025-Q1-K1-E10

Method BLASTX
NCBI GI g3402709
BLAST score 659
E value 3.0e-69
Match length 179
% identity 54

NCBI Description (AC004261) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10702

Contig ID 19453 1.R1040

5'-most EST LIB3139-116-P1-N1-F11

Seq. No. 10703

Contig ID 19455 1.R1040

5'-most EST LIB3050-025-Q1-K1-E6

Seq. No. 10704

Contig ID 19456 1.R1040

5'-most EST jC-gmf102220056b02a1

Method BLASTX
NCBI GI g3928543
BLAST score 1598
E value 1.0e-178
Match length 468



```
% identity
                  (AB016819) UDP-glucose glucosyltransferase [Arabidopsis
NCBI Description
                  thaliana]
Seq. No.
                  10705
Contig ID
                  19456 2.R1040
5'-most EST
                  hrw701061409.h1
                  BLASTX
Method
                  q3928543
NCBI GI
BLAST score
                  508
E value
                   2.0e-51
                   158
Match length
                   59
% identity
                  (AB016819) UDP-glucose glucosyltransferase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   10706
Contig ID
                   19456 3.R1040
                   jC-qmf102220146c05d1
5'-most EST
                   BLASTX
Method
                   g2827992
NCBI GI
BLAST score
                   309
                   3.0e-28
E value
                   70
Match length
                   77
% identity
                  (AF034743) UDP-glucuronosyltransferase [Pisum sativum]
NCBI Description
                   10707
Seq. No.
                   19456 4.R1040
Contig ID
5'-most EST
                   jC-gmro02800042g04d1
                   BLASTN
Method
NCBI GI
                   g2827991
BLAST score
                   107
                   5.0e-53
E value
                   239
Match length
                   86
% identity
                   Pisum sativum UDP-glucuronosyltransferase gene, complete
NCBI Description
                   10708
Seq. No.
                   19459 1.R1040
Contig ID
                   fC-gmst700605438f5
5'-most EST
                   BLASTX
Method
                   g3402751
NCBI GI
BLAST score
                   973
                   1.0e-105
E value
Match length
                   270
% identity
                   69
                  (AL031187) putative protein [Arabidopsis thaliana]
NCBI Description
                   10709
Seq. No.
Contig ID
                   19460 1.R1040
5'-most EST
                   jsh701063724.h1
```

1865

10710

19474_1.R1040 uaw700664864.h1

Seq. No. Contig ID

5'-most EST



Seq. No. 10711

Contig ID 19489_1.R1040 5'-most EST g4293317

Method BLASTX
NCBI GI g1730081
BLAST score 504
E value 9.0e-51
Match length 155

Match length 155 % identity 67

NCBI Description PROTEIN L1R18B (LLPR10.1B) >gi_1039336_emb_CAA56299_ (X79975) L1R18B [Lupinus luteus] >gi_2183277 (AF002278)

L1PR10.1B [Lupinus luteus]

Seq. No. 10712

Contig ID 19489 2.R1040 5'-most EST yz1700966962.h1

Method BLASTX
NCBI GI g1703042
BLAST score 367
E value 1.0e-34
Match length 157
% identity 42

NCBI Description ABA-RESPONSIVE PROTEIN ABR17 >gi_629631_pir__\$42649

abscisic acid-responsive protein ABR17 - garden pea >gi_20631_emb_CAA78829_ (Z15128) ABA-responsive protein

[Pisum sativum]

Seq. No. 10713

Contig ID 19489_3.R1040 5'-most EST seb700650946.h1

Method BLASTX
NCBI GI g1730080
BLAST score 163
E value 2.0e-11
Match length 58
% identity 57

NCBI Description PROTEIN L1R18A (LLPR10.1A) >gi_1039334_emb_CAA56298_ (X79974) L1R18A [Lupinus luteus] >gi_2183275 (AF002277)

LlPR10.1A [Lupinus luteus]

Seq. No. 10714

Contig ID 19490_1.R1040

5'-most EST LIB3109-029-Q1-K1-F7

Seq. No. 10715

Contig ID 19492 1.R1040

5'-most EST jC-gmf102220126c10a1

Seq. No. 10716

Contig ID 19492_2.R1040 5'-most EST hrw701058618.h1

Seq. No. 10717

Contig ID 19493 1.R1040

5'-most EST LIB3050-025-Q1-K1-B6

Method BLASTX



NCBI GI g3980387
BLAST score 379
E value 2.0e-36
Match length 112
% identity 62

NCBI Description (AC004561) putative glutathione S-transferase [Arabidopsis

thaliana]

Seq. No. 10718

Contig ID 19499 1.R1040

5'-most EST LIB3109-035-Q1-K3-F6

Method BLASTX
NCBI GI g3688181
BLAST score 361
E value 2.0e-34
Match length 130
% identity 58

NCBI Description (AL031804) putative protein (fragment) [Arabidopsis

thaliana]

Seq. No. 10719

Contig ID 19508 1.R1040

5'-most EST LIB3050-024-Q1-K1-H7

Seq. No. 10720

Contig ID 19509 1.R1040

5'-most EST LIB3092-050-Q1-K1-H9

Method BLASTX
NCBI GI g3881189
BLAST score 432
E value 2.0e-52
Match length 145

% identity 64

NCBI Description (Z99281) similar to ADP-ribosylation factor; cDNA EST

EMBL:C08179 comes from this gene; cDNA EST EMBL:C08337 comes from this gene; cDNA EST EMBL:C09829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4

Seq. No. 10721

Contig ID 19509 2.R1040

5'-most EST LIB3050-024-Q1-K1-H8

Method BLASTX
NCBI GI g3881189
BLAST score 337
E value 1.0e-31
Match length 84
% identity 71

NCBI Description (Z99281) similar to ADP-ribosylation factor; cDNA EST

EMBL:C08179 comes from this gene; cDNA EST EMBL:C08337 comes from this gene; cDNA EST EMBL:C09829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4

Seq. No. 10722

Contig ID 19515_1.R1040 5'-most EST rca701001392.h1

Method BLASTX NCBI GI g3914996



BLAST score 231 E value 3.0e-19 Match length 51 % identity 86

NCBI Description PHOSPHOSERINE AMINOTRANSFERASE PRECURSOR (PSAT) >qi 1665831 dbj BAA13640 (D88541) phosphoserine

aminotransferase [Arabidopsis thaliana]

>gi_2804260_dbj_BAA24441_ (AB010408) phosphoserine

aminotransferase [Arabidopsis thaliana]

>gi_3367581_emb_CAA20033_ (AL031135) phosphoserine

aminotransferase [Arabidopsis thaliana]

Seq. No. 10723

Contig ID 19515 2.R1040

5'-most EST LIB3093-016-Q1-K1-C11

Method BLASTN
NCBI GI g2804259
BLAST score 150
E value 2.0e-78
Match length 446
% identity 83

NCBI Description Arabidopsis thaliana DNA for phosphoserine

aminotransferase, complete cds

Seq. No. 10724

Contig ID 19515_3.R1040 5'-most EST jsh701065269.h1

Method BLASTN
NCBI GI g2804259
BLAST score 93
E value 1.0e-44
Match length 189
% identity 88

NCBI Description Arabidopsis thaliana DNA for phosphoserine

aminotransferase, complete cds

Seq. No. 10725

Contig ID 19516_1.R1040

5'-most EST jC-gmf102220050c05a1

Seq. No.

10726

Contig ID 19517 1.R1040

5'-most EST LIB3050-024-Q1-K1-F11

Method BLASTX
NCBI GI g2832658
BLAST score 244
E value 2.0e-20
Match length 94
% identity 51

NCBI Description (AL021710) putative protein [Arabidopsis thaliana]

Seq. No. 10727

Contig ID 19517_2.R1040 5'-most EST zzp700834035.h1

Method BLASTX
NCBI GI g2832658
BLAST score 193



```
4.0e-15
E value
Match length
                   58
                   57
% identity
```

(AL021710) putative protein [Arabidopsis thaliana] NCBI Description

10728 Seq. No.

19518 1.R1040 Contig ID

LIB3050-024-Q1-K1-F12 5'-most EST

10729 Seq. No.

19520 1.R1040 Contig ID

LIB3050-024-Q1-K1-F4 5'-most EST

Method BLASTX g417148 NCBI GI BLAST score 379 E value 3.0e-39 138 Match length % identity 61

PROBABLE GLUTATHIONE S-TRANSFERASE (HEAT SHOCK PROTEIN 26A) NCBI Description

(G2-4) >gi 99912_pir A33654 heat shock protein 26A soybean $>g\overline{1}_{169981}$ ($\overline{M2}_{0363}$) Gmhsp26-A [Glycine max]

10730 Seq. No.

19521 1.R1040 Contig ID wrg700788594.h1 5'-most EST

Method BLASTX q4335715 NCBI GI BLAST score 586 E value 9.0e-71 291 Match length % identity 51

(AC006248) putative immediate-early salicylate-induced NCBI Description

glucosyltransferase [Arabidopsis thaliana]

10731 Seq. No.

19528 1.R1040 Contig ID uC-gmropic010a04b1 5'-most EST

Seq. No. 10732

19530 1.R1040 Contig ID 5'-most EST pxt700945228.h1

Method BLASTX g4510375 NCBI GI 207 BLAST score 3.0e-16 E value 123 Match length % identity

(AC007017) putative homeotic protein BEL1 [Arabidopsis NCBI Description

thaliana]

Seq. No. 10733

Contig ID 19533 1.R1040 bth700843865.h1 5'-most EST

10734 Seq. No.

Contig ID 19534 1.R1040

5'-most EST LIB3050-024-Q1-K1-D6



Contig ID 19536_1.R1040

5'-most EST LIB3050-024-Q1-K1-D8

Seq. No. 10736

Contig ID 19540_1.R1040

5'-most EST LIB3050-024-Q1-K1-E11

Seq. No. 10737

Contig ID 19540 2.R1040

5'-most EST jC-gmf102220052c02d1

Seq. No. 10738

Contig ID 19541_1.R1040 5'-most EST zzp700832101.h1

Method BLASTX
NCBI GI g2651310
BLAST score 515
E value 4.0e-52
Match length 236
% identity 43

NCBI Description (AC002336) putative PTR2-B peptide transporter [Arabidopsis

thaliana]

Seq. No. 10739

Contig ID 19543 1.R1040

5'-most EST LIB3074-028-Q1-K1-C4

Seq. No. 10740

Contig ID 19544_1.R1040 5'-most EST fde700870563.h1

Method BLASTX
NCBI GI 94455225
BLAST score 757
E value 5.0e-99
Match length 198
% identity 84

NCBI Description (AL035440) putative protein [Arabidopsis thaliana]

Seq. No.

10741

Contig ID 19545 1.R1040

5'-most EST uC-gmrominsoy048b10b1

10742

Method BLASTN
NCBI GI g1142620
BLAST score 105
E value 6.0e-52
Match length 354
% identity 89

NCBI Description Phaseolus vulgaris phaseolin G-box binding protein PG2

(PG2) mRNA, partial cds

Seq. No.

Contig ID 19546 1.R1040 5'-most EST pmv700893028.h1

Method BLASTX NCBI GI g3080420



```
BLAST score 534
E value 2.0e-54
Match length 144
% identity 67
```

NCBI Description (AL022604) putative sugar transporter protein [Arabidopsis

thaliana]

Seq. No. 10743

Contig ID 19548 1.R1040

5'-most EST LIB3050-024-Q1-K1-B8

Seq. No. 10744

Contig ID 19553_1.R1040 5'-most EST pxt700944431.h1

Seq. No. 10745

Contig ID 19565_1.R1040 5'-most EST rca701002450.h1

Method BLASTN
NCBI GI 9719289
BLAST score 33
E value 7.0e-09
Match length 85
% identity 85

NCBI Description Arabidopsis thaliana SABRE gene, exon 14 and complete cds

Seq. No. 10746

Contig ID 19568 1.R1040

5'-most EST LIB3107-013-Q1-K1-A2

Method BLASTN
NCBI GI g945059
BLAST score 536
E value 0.0e+00
Match length 907
% identity 90

NCBI Description Vigna unguiculata aminoimidazole ribonucleotide (AIRS)

synthetase (pur5) mRNA, complete cds

Seq. No. 10747

Contig ID 19568 2.R1040

5'-most EST LIB3170-062-Q1-K1-F10

Method BLASTX
NCBI GI g1709925
BLAST score 187
E value 4.0e-14
Match length 64
% identity 64

NCBI Description PHOSPHORIBOSYLFORMYLGLYCINAMIDINE CYCLO-LIGASE PRECURSOR

(AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZOLE SYNTHETASE) (AIR SYNTHASE) >gi_945060 (U30895) aminoimidazole ribonucleotide

(AIRS) synthetase [Vigna unguiculata]

Seq. No. 10748

Contig ID 19569 1.R1040 5'-most EST leu701156202.h1

Method BLASTX NCBI GI g421989



```
BLAST score
                   999
                   1.0e-108
E value
Match length
                   387
                   55
% identity
                   serpin - barley >gi_19071_emb_CAA78822_ (Z15116) protein zx
NCBI Description
                   [Hordeum vulgare] >\overline{gi}_444\overline{7}78_prf_1908\overline{2}13A protein Zx
                   [Hordeum vulgare]
                   10749
Seq. No.
                   19569 3.R1040
Contig ID
5'-most EST
                   wrg700787582.h1
                   BLASTX
Method
                   g421989
NCBI GI
BLAST score
                   150
                   1.0e-09
E value
                   59
Match length
                   51
% identity
                   serpin - barley >gi_19071_emb_CAA78822_ (Z15116) protein zx
NCBI Description
                    [Hordeum vulgare] >\overline{gi}_444\overline{778}_prf 1908\overline{2}13A protein Zx
                    [Hordeum vulgare]
                   10750
Seq. No.
                   19582 1.R1040
Contig ID
                   gbt700547938.h1
5'-most EST
                   BLASTX
Method
                   q4544409
NCBI GI
                                                                             ege.
BLAST score
                   343
                   5.0e-32
E value
Match length
                   103
                   59
% identity
                    (AC006955) putative transcription factor [Arabidopsis
NCBI Description
                    thaliana]
                    10751
Seq. No.
                    19582 2.R1040
Contig ID
5'-most EST
                   LIB3050-023-Q1-K1-G4
                   BLASTX
Method
NCBI GI
                    g4544409
BLAST score
                    224
                    2.0e-18
E value
                    69
Match length
% identity
                    58
                   (AC006955) putative transcription factor [Arabidopsis
NCBI Description
                    thaliana]
                    10752
Seq. No.
                    19584_1.R1040
Contig ID
                    LIB3050-023-Q1-K1-G6
5'-most EST
                    BLASTX
Method
NCBI GI
                    g4567203
BLAST score
                    277
E value
                    1.0e-24
                    63
Match length
```

(AC007168) putative beta-hydroxyacyl-ACP dehydratase

84

[Arabidopsis thaliana]

% identity

NCBI Description



10753 Seq. No. 19586_1.R1040 Contig ID leu701150647.h1 5'-most EST Seq. No. 10754 19586 2.R1040 Contig ID jC-qmle01810011h12a1 5'-most EST Seq. No. 10755 Contig ID 19590 1.R1040 LIB3050-023-Q1-K1-H12 5'-most EST Seq. No. 10756 Contig ID 19591 1.R1040 5'-most EST LIB3050-023-Q1-K1-H2 10757 Seq. No. Contig ID 19593 1.R1040 LIB3139-069-P1-N1-G7 5'-most EST BLASTX Method NCBI GI q1838976 BLAST score 289 E value 5.0e-50 Match length 243 49 % identity (X73635) vsf-1 [Lycopersicon esculentum] NCBI Description 10758 Seq. No. 19593 2.R1040 Contig ID smc700750091.h1 5'-most EST Seq. No. 10759 19595 1.R1040 Contig ID $r1r70\overline{0}901422.h1$ 5'-most EST 10760 Seq. No. 19596 1.R1040 Contig ID zsq701128870.h1 5'-most EST BLASTX Method g2129915 NCBI GI 493 BLAST score 2.0e-49 E value 148 Match length 66 % identity ferredoxin precursor - sweet orange NCBI Description >gi_1360725_emb_CAA87068_ (Z46944) non-photosynthetic ferredoxin [Citrus sinensis] 10761 Seq. No. 19599 1.R1040 Contig ID 5'-most EST fC-gmse700660496f5

 Seq. No.
 10761

 Contig ID
 19599_1.R1040

 5'-most EST
 fC-gmse700660496f5

 Method
 BLASTX

 NCBI GI
 g3193332

 BLAST score
 304

 E value
 1.0e-27

Match length 164 % identity 51



(AF069299) similar to Arabidopsis AT-hook protein 1 NCBI Description (GB:AJ222585) [Arabidopsis thaliana] Seq. No. 10762 Contig ID 19607 1.R1040 uC-gmrominsoy316d05b1 5'-most EST Method BLASTX q1002796 NCBI GI BLAST score 602 1.0e-62 E value 140 Match length 78 % identity (U33915) Cpm10 [Craterostigma plantagineum] NCBI Description Seq. No. 10763 19607 2.R1040 Contig ID 5'-most EST LIB3050-023-Q1-K1-F5 Method BLASTX g1002796 NCBI GI BLAST score 291 E value 2.0e-38 Match length 157 % identity 57 (U33915) Cpm10 [Craterostigma plantagineum] NCBI Description Seq. No. 10764 19611 1.R1040 Contig ID 5'-most EST LIB3139-095-P1-N1-B11 Method BLASTX q4006829 NCBI GI BLAST score 1516 1.0e-169 E value 429 Match length % identity 70 (AC005970) putative protein kinase [Arabidopsis thaliana] NCBI Description 10765 Seq. No. 19612 1.R1040 Contig ID 5'-most EST LIB3106-065-P1-K1-F12 Method BLASTX NCBI GI g4490702 BLAST score 213 7.0e-17 E value Match length 63 % identity 65 (AL035680) hypothetical protein [Arabidopsis thaliana] NCBI Description 10766 Seq. No.

Contig ID 19615_1.R1040

5'-most EST LIB3050-023-Q1-K1-E8

Method BLASTX
NCBI GI g2759999
BLAST score 466
E value 2.0e-46
Match length 179
% identity 49

NCBI Description (AJ003141) peroxidase [Hordeum vulgare]



Contig ID 19619_1.R1040

5'-most EST uC-gmflminsoy042g05b1

Method BLASTX
NCBI GI g2244792
BLAST score 2456
E value 0.0e+00
Match length 601
% identity 76

NCBI Description (Z97336) ankyrin homolog [Arabidopsis thaliana]

Seq. No.

10768

Contig ID 19619_2.R1040

5'-most EST uC-gmrominsoy061a01b1

Seq. No.

10769

Contig ID 19619 4.R1040

5'-most EST uC-gmrominsoy044a02b1

Seq. No. 10770

Contig ID 19622_1.R1040

5'-most EST LIB3106-091-Q1-K1-G4

Method BLASTX
NCBI GI g1749676
BLAST score 361
E value 1.0e-33
Match length 247
% identity 37

NCBI Description (D89234) similar to Saccharomyces cerevisiae ORF YGR205W,

EMBL Accession Number Z72990 [Schizosaccharomyces pombe]

Seq. No. 10771

Contig ID 19622_2.R1040 5'-most EST fde700874652.h1

Seq. No. 10772

Contig ID 19625_1.R1040

5'-most EST LIB3109-033-Q1-K1-F3

Seq. No. 10773

Contig ID 19625_2.R1040

5'-most EST LIB3139-046-P1-N1-E1

Method BLASTX
NCBI GI g4544409
BLAST score 716
E value 1.0e-75
Match length 186
% identity 70

NCBI Description (AC006955) putative transcription factor [Arabidopsis

thaliana]

Seq. No. 10774

Contig ID 19625 3.R1040

5'-most EST jC-gmst02400009h01a1

Seq. No. 10775



```
19625 4.R1040
Contig ID
                  uC-gmrominsoy231b03b1
5'-most EST
                  BLASTX
Method
NCBI GI
                   q4544409
                   259
BLAST score
                   2.0e-22
E value
                  88
Match length
                   53
% identity
                  (AC006955) putative transcription factor [Arabidopsis
NCBI Description
                   thaliana]
                   10776
Seq. No.
                   19626 1.R1040
Contig ID
                   ejt700606274.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1730560
BLAST score
                   2686
                   0.0e + 00
E value
                   557
Match length
                   91
% identity
                   ALPHA-GLUCAN PHOSPHORYLASE, H ISOZYME (STARCH PHOSPHORYLASE
NCBI Description
                   H) >gi_510932_emb_CAA84494_ (Z35117) alpha 1,4-glucan
                   phosphorylase type H [Vicia faba]
                   10777
Seq. No.
                   19627 1.R1040
Contig ID
                   LIB3050-023-Q1-K1-E3
5'-most EST
                   10778
Seq. No.
                   19633 1.R1040
Contig ID
                   crh70\overline{0}851363.h1
5'-most EST
Method
                   BLASTX
                   q4455251
NCBI GI
                   625
BLAST score
                   8.0e-77
E value
                   206
Match length
                   75
% identity
                   (AL035523) magnesium-protoporphyrin IX
NCBI Description
                   methyltransferase-like protein [Arabidopsis thaliana]
                   10779
Seq. No.
                   19640 1.R1040
 Contig ID
                   LIB3050-023-Q1-K1-C7
 5'-most EST
                   10780
 Seq. No.
```

19644 1.R1040 Contig ID taw700659052.hl 5'-most EST

BLASTX Method g4093157 NCBI GI 734 BLAST score 2.0e-77 E value Match length 274 % identity

(AF087936) phytochrome-associated protein 2 [Arabidopsis NCBI Description thaliana]

10781 Seq. No.



Contig ID 19644_2.R1040 5'-most EST taw700657766.h1

Method BLASTX
NCBI GI g4093157
BLAST score 412
E value 3.0e-40
Match length 118
% identity 77

NCBI Description (AF087936) phytochrome-associated protein 2 [Arabidopsis

thaliana]

Seq. No. 10782

Contig ID 19648 1.R1040

5'-most EST LIB $30\overline{5}0-023-Q1-K1-A1$

Method BLASTX
NCBI GI g3157933
BLAST score 343
E value 6.0e-33
Match length 106
% identity 61

NCBI Description (AC002131) Contains similarity to box helicases gb_U29097

from C. elegans and to the ENBP1 gene product gb_X95995

from Vicia sativa. [Arabidopsis thaliana]

Seq. No. 10783

Contig ID 19648_2.R1040 5'-most EST fde700876682.h1

Method BLASTX
NCBI GI g2781345
BLAST score 427
E value 4.0e-42
Match length 96
% identity 80

NCBI Description (AC003113) F2401.2 [Arabidopsis thaliana]

Seq. No. 10784

Contig ID 19652_1.R1040 5'-most EST taw700660047.h1

Seq. No. 10785

Contig ID 19653_1.R1040

5'-most EST LIB3107-030-Q1-K1-B6

Method BLASTX
NCBI GI g4220527
BLAST score 896
E value 2.0e-96
Match length 315
% identity 56

NCBI Description (AL035356) putative protein [Arabidopsis thaliana]

Seq. No. 10786

Contig ID 19653_2.R1040 5'-most EST pxt700941158.h1

Seq. No. 10787

Contig ID 19654 1.R1040

5'-most EST LIB3093-005-Q1-K1-E7



```
10788
Seq. No.
                  19654 2.R1040
Contig ID
                  uC-gmrominsoy306b07b1
5'-most EST
                  10789
Seq. No.
                  19662 1.R1040
Contig ID
5'-most EST
                  LIB3139-065-P1-N1-H3
Method
                  BLASTX
                  g3297817
NCBI GI
                  153
BLAST score
                  1.0e-09
E value
                  79
Match length
                  39
% identity
                  (AL031032) putative protein [Arabidopsis thaliana]
NCBI Description
                  10790
Seq. No.
                  19662 2.R1040
Contig ID
5'-most EST
                  LIB3050-014-Q1-E1-B10
Seq. No.
                  10791
                  19662 3.R1040
Contig ID
                  uC-gmrominsoy189h09b1
5'-most EST
                  10792
Seq. No.
Contig ID
                  19663 1.R1040
5'-most EST
                  LIB3051-036-Q1-K1-H5
Method
                  BLASTX
                   q3004556
NCBI GI
BLAST score
                   355
                   4.0e-33
E value
                  236
Match length
                   38
% identity
                  (AC003673) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10793
Seq. No.
                   19665 1.R1040
Contig ID
5'-most EST
                   LIB3109-006-Q1-K1-E3
                   BLASTX
Method
NCBI GI
                   g2649424
BLAST score
                   157
E value
                   2.0e-10
Match length
                   100
                   39
% identity
                  (AE001023) A. fulgidus predicted coding region AF1178
NCBI Description
                   [Archaeoglobus fulgidus]
                   10794
Seq. No.
Contig ID
                   19667 1.R1040
                   uC-gmronoir031d10b1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q4263507
BLAST score
                   735
```

E value 2.0e-77 Match length 588 % identity 30

NCBI Description (AC004044) hypothetical protein [Arabidopsis thaliana]



Contig ID 19667_2.R1040 5'-most EST uC-gmropic101b12b1

Seq. No. 10796

Contig ID 19669 1.R1040

5'-most EST LIB3050-022-Q1-K1-G7

Seq. No. 10797

Contig ID 19671 1.R1040

5'-most EST LIB3074-036-Q1-K1-G12

Seq. No. 10798

Contig ID 19671_2.R1040 5'-most EST pxt700944051.h1

Seq. No. 10799

Contig ID 19680 1.R1040

5'-most EST LIB3107-063-Q1-K1-E11

Seq. No. 10800

Contig ID 19682 1.R1040

5'-most EST LIB3073-019-Q1-K1-B7

Method BLASTX
NCBI GI g2245001
BLAST score 459
E value 1.0e-45
Match length 101
% identity 85

NCBI Description (Z97341) similarity to NADH dehydrogenase (ubiquinone)

[Arabidopsis thaliana]

Seq. No. 10801

Contig ID 19685_1.R1040 5'-most EST wrg700791916.h1

Method BLASTX
NCBI GI g3617741
BLAST score 906
E value 5.0e-98
Match length 213
% identity 79

NCBI Description (AC005687) L3 cytoplasmic ribosomal protein [Arabidopsis

thaliana]

Seq. No. 10802

Contig ID 19686_1.R1040

5'-most EST uC-gmrominsoy056g12b1

Method BLASTX
NCBI GI g3128172
BLAST score 251
E value 3.0e-21
Match length 88
% identity 48

NCBI Description (AC004521) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10803



19686 2.R1040 Contig ID 5'-most EST uC-gmropic010c03b1 BLASTX Method g3128172 NCBI GI BLAST score 695 E value 2.0e-73 188 Match length 65 % identity (AC004521) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 10804 19689 1.R1040 Contig ID zpv700756991.h1 5'-most EST BLASTX Method g4335763 NCBI GI BLAST score 404 E value 3.0e-39 Match length 159 % identity (AC006284) unknown protein [Arabidopsis thaliana] NCBI Description Seq. No. 10805 19689 2.R1040 Contig ID LIB3139-082-P1-N1-D4 5'-most EST BLASTX Method NCBI GI q4335763 BLAST score 222 E value 4.0e-18 Match length 96 % identity 42 (AC006284) unknown protein [Arabidopsis thaliana] NCBI Description 10806 Seq. No. 19693 1.R1040 Contig ID zsq701118870.hl 5'-most EST Method BLASTX g3249066 NCBI GI BLAST score 284 3.0e-25 E value Match length 117 56 % identity (AC004473) Similar to S. cerevisiae SIK1P protein NCBI Description gb 984964. ESTs gb F15433 and gb AA395158 come from this gene. [Arabidopsis thaliana] 10807 Seq. No. 19696_1.R1040 Contig ID 5'-most EST LIB3050-025-Q1-K1-C12 Seq. No. 10808 Contig ID 19697 1.R1040

 $ncj70\overline{0}975670.h1$ 5'-most EST

Method BLASTX g1420936 NCBI GI 1820 BLAST score E value 0.0e + 00Match length 407



% identity 8

NCBI Description (U61396) Vigna unguiculata aspartic proteinase mRNA,

complete cds. [Vigna unguiculata]

Seq. No. 10809

Contig ID 19703_1.R1040

5'-most EST LIB3050-022-Q1-K1-E1

Method BLASTX
NCBI GI g2583123
BLAST score 205
E value 4.0e-16
Match length 47
% identity 77

NCBI Description (AC002387) putative nucleotide sugar epimerase [Arabidopsis

thaliana]

Seq. No. 10810

Contig ID 19706_1.R1040

5'-most EST LIB3109-005-Q1-K1-B6

Method BLASTX
NCBI GI g3023625
BLAST score 324
E value 2.0e-29
Match length 148
% identity 44

NCBI Description POLYPEPTIDE DEFORMYLASE (PDF) (FORMYLMETHIONINE

DEFORMYLASE) >gi_1772503_emb_CAA71353_ (Y10305) polypeptide

deformylase [Calothrix PCC7601]

Seq. No. 10811

Contig ID 19713_1.R1040

5'-most EST jC-gmf102220073d11a1

Seq. No. 10812

Contig ID 19713 2.R1040

5'-most EST LIB3050-022-Q1-K1-E8

Seq. No. 10813

Contig ID 19714 1.R1040

5'-most EST LIB3050-022-Q1-K1-D1

Method BLASTX
NCBI GI g4512659
BLAST score 369
E value 2.0e-35
Match length 99
% identity 73

NCBI Description (AC006931) putative protein kinase [Arabidopsis thaliana]

>gi_4544465_gb_AAD22372.1_AC006580_4 (AC006580) putative

protein kinase [Arabidopsis thaliana]

Seq. No. 10814

Contig ID 19715_1.R1040

5'-most EST LIB3050-019-Q1-K1-H7

Seq. No. 10815

Contig ID 19720_1.R1040

5'-most EST LIB3050-022-Q1-K1-C3



```
10816
Seq. No.
                  19724 1.R1040
Contig ID
                  fua701038254.hl
5'-most EST
                  BLASTX
Method
                  g2245060
NCBI GI
BLAST score
                  686
                  5.0e-72
E value
                  204
Match length
                  61
% identity
NCBI Description (Z97342) allergen homolog [Arabidopsis thaliana]
                  10817
Seq. No.
                  19728 1.R1040
Contig ID
                  LIB3050-022-Q1-K1-D10
5'-most EST
                  BLASTX
Method
                  g3386620
NCBI GI
                   592
BLAST score
                   2.0e-61
E value
                   134
Match length
% identity
                   46
                  (AC004665) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10818
Seq. No.
                   19729 1.R1040
Contig ID
                   uC-gmrominsoy048d04b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3335376
                   1170
BLAST score
                   1.0e-131
E value
Match length
                   313
                   75
% identity
                   (AC003028) putative ammonium transporter [Arabidopsis
NCBI Description
                   thaliana]
                   10819
Seq. No.
                   19729 2.R1040
Contig ID
5'-most EST
                   uC-gmropic115g04b1
                   BLASTX
Method
                   g3335376
NCBI GI
BLAST score
                   271
                   8.0e-24
E value
Match length
                   58
% identity
                   90
                   (AC003028) putative ammonium transporter [Arabidopsis
NCBI Description
                   thaliana]
                   10820
Seq. No.
                   19733 1.R1040
Contig ID
                   LIB3050-022-Q1-K1-A5
5'-most EST
Method
                   BLASTX
                   q3643608
NCBI GI
BLAST score
                   372
                   9.0e-36
E value
                   129
Match length
```

1882

(AC005395) hypothetical protein [Arabidopsis thaliana]

61

% identity

NCBI Description

```
10821
Seq. No.
                  19737 1.R1040
Contig ID
                  LIB3050-022-Q1-K1-A9
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3461846
                  312
BLAST score
                  2.0e-28
E value
                  91
Match length
                   33
% identity
                  (AC005315) putative zinc-finger protein [Arabidopsis
NCBI Description
                  thaliana]
                   10822
Seq. No.
                   19739 1.R1040
Contig ID
                   fde700875824.h1
5'-most EST
                   BLASTX
Method
                   g2829910
NCBI GI
BLAST score
                   793
                   9.0e-85
E value
Match length
                   166
                   19
% identity
NCBI Description (AC002291) Unknown protein, contains regulator of
                   chromosome condensation motifs [Arabidopsis thaliana]
                   10823
Seq. No.
                   19739 2.R1040
Contig ID
                   r1r70\overline{0}896417.h1
5'-most EST
                   BLASTX
Method
                   q2829910
NCBI GI
                   757
BLAST score
                   1.0e-80
E value
Match length
                   191
% identity
                   (AC002291) Unknown protein, contains regulator of
NCBI Description
                   chromosome condensation motifs [Arabidopsis thaliana]
Seq. No.
                   10824
                   19744 1.R1040-
Contig ID
                   LIB3050-022-Q1-K1-B4
5'-most EST
                   10825
Seq. No.
Contig ID
                   19746 1.R1040
```

pcp700989286.h1 5'-most EST

10826 Seq. No.

19756 1.R1040 Contig ID

uC-gmrominsoy183d12b1 5'-most EST

BLASTX Method g4510345 NCBI GI 460 BLAST score E value 2.0e-45 160 Match length 54 % identity

(AC006921) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No. 10827



19756 2.R1040 Contig ID LIB3170-078-Q1-K1-E5 5'-most EST BLASTX Method g4510345 NCBI GI BLAST score 244 2.0e-20 E value 93 Match length 77 % identity (AC006921) unknown protein [Arabidopsis thaliana] NCBI Description 10828 Seq. No. 19756 3.R1040 Contig ID LIB3139-085-P1-N1-H5 5'-most EST BLASTX Method NCBI GI g4510345 BLAST score 232 3.0e-19 E value Match length 57 70 % identity (AC006921) unknown protein [Arabidopsis thaliana] NCBI Description 10829 Seq. No. 19756 4.R1040 Contig ID fua701039811.hl 5'-most EST Method BLASTX NCBI GI g4510345 BLAST score 170 E value 3.0e-12Match length 33 % identity 88 (AC006921) unknown protein [Arabidopsis thaliana] NCBI Description 10830 Seq. No. 19760 1.R1040 Contig ID hyd700726331.h1 5'-most EST Method BLASTX NCBI GI g2827141 BLAST score 1633 0.0e+00E value 371 Match length % identity 81 (AF027173) cellulose synthase catalytic subunit NCBI Description [Arabidopsis thaliana] 10831 Seq. No. 19764 1.R1040 Contig ID LIB3109-001-Q1-K4-E2 5'-most EST 10832 Seq. No. 19767_1.R1040 Contig ID 5'-most EST LIB3093-018-Q1-K2-D3 Method BLASTX

NCBI GI g1703143 BLAST score 568 3.0e-58 E value Match length 164 % identity 65



NCBI Description ACTIN-LIKE PROTEIN 3 >gi_881635 (U29610) Arp3 [Acanthamoeba castellanii]

Seq. No. 10833

Contig ID 19767_2.R1040

5'-most EST jC-gmst02400023b08a1

Method BLASTX
NCBI GI g1168329
BLAST score 498
E value 3.0e-50
Match length 175
% identity 54

NCBI Description ACTIN-LIKE PROTEIN 3 (ACTIN-LIKE PROTEIN 66B) (ACTIN-2)

>gi_558568_emb_CAA50674_ (X71789) actin related protein
[Drosophila melanogaster] >gi_1096138_prf__2111232A

actin-related protein [Drosophila melanogaster]

Seq. No. 10834

Contig ID 19768_1.R1040

5'-most EST LIB3107-053-Q1-K1-A5

Method BLASTX
NCBI GI g2760834
BLAST score 948
E value 1.0e-103
Match length 278
% identity 70

NCBI Description (AC003105) putative nitrate transporter [Arabidopsis

thaliana]

Seq. No. 10835

Contig ID 19768_2.R1040 5'-most EST kl1701205306.h1

Method BLASTX
NCBI GI g3912921
BLAST score 148
E value 3.0e-09
Match length 59
% identity 46

NCBI Description (AF001308) G/T DNA mismatch repair enzyme [Arabidopsis

thaliana]

Seq. No. 10836

Contig ID 19768_3.R1040 5'-most EST taw700657702.h1

Seq. No. 10837

Contig ID 19769 1.R1040

5'-most EST LIB30 $\overline{5}$ 0-021-Q1-K1-C2

Method BLASTX
NCBI GI g4454567
BLAST score 235
E value 2.0e-19
Match length 123
% identity 45

NCBI Description (AF128407) lipase homolog [Arabidopsis thaliana]

Seq. No. 10838



```
19774 1.R1040
Contig ID
                  LIB30\overline{5}0-021-Q1-K1-C9
5'-most EST
                  BLASTX
Method
                  g2147328
NCBI GI
BLAST score
                  202
E value
                  5.0e-16
                   90
Match length
                   52
% identity
NCBI Description dioscorin class B - Dioscorea cayenensis (fragment)
                   10839
Seq. No.
                   19775 1.R1040
Contig ID
                  LIB3109-045-Q1-K1-D6
5'-most EST
                   BLASTX
Method
                   g2160184
NCBI GI
BLAST score
                   143
E value
                   7.0e-09
                   107
Match length
                   41
% identity
                   (AC000132) ESTs gb H37208,gb_H36853 come from this gene.
NCBI Description
                   [Arabidopsis thaliana]
                   10840
Seq. No.
                   19790 1.R1040
Contig ID
                   LIB3107-006-Q1-K1-B5
5'-most EST
                   BLASTX
Method
                   g2501448
NCBI GI
                   450
BLAST score
                   9.0e-45
E value
Match length
                   93
% identity
                   UBIQUITIN-LIKE PROTEIN SMT3 >gi_1707372_emb_CAA67923
NCBI Description
                   (X99609) ubiquitin-like protein [Arabidopsis thaliana]
                   10841
Seq. No.
                   19791 1.R1040
Contig ID
5'-most EST
                   g5677872
                   10842
Seq. No.
                   19791 2.R1040
Contig ID
                   LIB3050-021-Q1-K1-E12
 5'-most EST
 Seq. No.
                   10843
                   19794 1.R1040
 Contig ID
                   LIB3139-040-P1-N1-A3
 5'-most EST
                   BLASTX
 Method
                   g4558591
 NCBI GI
                   588
 BLAST score
                   2.0e-60
 E value
                   178
 Match length
 % identity
                   (AC006555) putative beta-1,3-glucanase [Arabidopsis
 NCBI Description
                   thaliana]
                   10844
 Seq. No.
```

1886

19794 2.R1040

sat701004908.h1

Contig ID

5'-most EST

Contig ID 5'-most EST

10850

19797 1.R1040

LIB3050-020-Q1-K1-H6



```
Method
                  BLASTX
NCBI GI
                  g4455206
BLAST score
                  1230
                  1.0e-136
E value
Match length
                  324
% identity
                  72
                   (AL035440) putative beta-1, 3-glucanase [Arabidopsis
NCBI Description
                  10845
Seq. No.
                  19794 3.R1040
Contig ID
                  LIB3092-045-Q1-K1-A4
5'-most EST
                  BLASTX
Method
                   q4455206
NCBI GI
BLAST score
                   351
E value
                   3.0e-33
Match length
                   99
% identity
                   (AL035440) putative beta-1, 3-glucanase [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   10846
                   19794 4.R1040
Contig ID
                   LIB3092-030-Q1-K1-D5
5'-most EST
                   BLASTX
Method
                   q4455206
NCBI GI
BLAST score
                   251
                   2.0e-21
E value
                   77
Match length
% identity
                   61
                   (AL035440) putative beta-1, 3-glucanase [Arabidopsis
NCBI Description
                   thaliana]
                   10847
Seq. No.
                   19794 5.R1040
Contig ID
                   jC-gmro02910043a11d1
5'-most EST
Method
                   BLASTX
                   q4558591
NCBI GI
BLAST score
                   242
                   2.0e-20
E value
Match length
                   62
                   71
% identity
                   (AC006555) putative beta-1,3-glucanase [Arabidopsis
NCBI Description
                   thaliana]
                   10848
Seq. No.
                   19794_6.R1040
Contig ID
                   jC-gmro02800026b05d1
5'-most EST
                   10849
Seq. No.
Contig ID
                   19795 1.R1040
                   LIB3050-020-Q1-K1-H4
5'-most EST
```



10851 Seq. No. 19801 1.R1040 Contig ID hyd700727614.hl 5'-most EST 10852 Seq. No. Contig ID 19803 1.R1040 pmv700888283.hl 5'-most EST 10853 Seq. No. 19804 1.R1040 Contig ID g4283486 5'-most EST BLASTX Method g2104816 NCBI GI 220 BLAST score 2.0e-17 E value 174 Match length 28 % identity NCBI Description (Y13092) alpha-adaptin [Drosophila melanogaster] 10854 Seq. No. 19807 1.R1040 Contig ID LIB3050-021-Q1-K1-A6 5'-most EST BLASTX Method g2342682 NCBI GI 290 BLAST score 6.0e-26 E value 76 Match length 72 % identity NCBI Description (AC000106) Contains similarity to Rattus AMP-activated protein kinase (gb_X95577). [Arabidopsis thaliana] 10855 Seq. No. Contig ID 19820 1.R1040 LIB3050-015-Q1-E1-D10 5'-most EST BLASTX Method g1871577 NCBI GI BLAST score 433 E value 1.0e-42 Match length 158 53 % identity (Y11553) putative 21kD protein precursor [Medicago sativa] NCBI Description Seq. No. 10856 19824 1.R1040 Contig ID LIB3050-020-Q1-K1-F2 5'-most EST Method BLASTX g4508069 NCBI GI BLAST score 385

5.0e-37 E value 149 Match length 48 % identity

(AC005882) 12246 [Arabidopsis thaliana] NCBI Description

10857 Seq. No.

19838 1.R1040 Contig ID

uC-gmrominsoy244f07b1 5'-most EST

BLASTX Method



g3785983 NCBI GI BLAST score 320 3.0e-29 E value Match length 153 % identity 43 NCBI Description (AC005560) hypothetical protein [Arabidopsis thaliana] 10858 Seq. No. 19838 2.R1040 Contig ID LIB3073-017-Q1-K1-E3 5'-most EST BLASTX Method g3785983 NCBI GI 253 BLAST score 1.0e-21 E value Match length 119 % identity 43 NCBI Description (AC005560) hypothetical protein [Arabidopsis thaliana] Seq. No. 10859 19838 3.R1040 Contig ID leu701149924.h1 5'-most EST 10860 Seq. No. 19842 1.R1040 Contig ID LIB3092-062-Q1-K1-C5 5'-most EST Method BLASTX g1931638 NCBI GI BLAST score 706 2.0e-74 E value Match length 277 58 % identity (U95973) transcription factor RUSH-lalpha isolog NCBI Description [Arabidopsis thaliana] 10861 Seq. No. Contig ID 19844 1.R1040 5'-most EST g4292527

Seq. No.

10862

19846 1.R1040 Contig ID

5'-most EST LIB3050-020-Q1-K1-C6

Seq. No.

10863

19852 1.R1040 Contig ID

5'-most EST LIB3051-107-Q1-K1-E10

Method BLASTN NCBI GI g1209887 BLAST score 279 E value 1.0e-155 703 Match length % identity

NCBI Description Glycine max nonsymbiotic hemoglobin gene, complete cds

Seq. No.

10864

19852 2.R1040 Contig ID

5'-most EST LIB3170-074-Q1-K1-E2

Method BLASTN



NCBI GI g1209887 BLAST score 404 E value 0.0e+00 Match length 420 % identity 99

NCBI Description Glycine max nonsymbiotic hemoglobin gene, complete cds

Seq. No. 10865

Contig ID 19854_1.R1040 5'-most EST epx701103941.h1

Method BLASTX
NCBI GI g2832717
BLAST score 325
E value 4.0e-30
Match length 99
% identity 67

NCBI Description (AJ003114) alkaline/neutral invertase [Lolium temulentum]

Seq. No. 10866

Contig ID 19855_1.R1040

5'-most EST LIB3050-025-Q1-K1-G12

Method BLASTX
NCBI GI g2982456
BLAST score 328
E value 3.0e-30
Match length 101
% identity 69

NCBI Description (AL022223) putative protein [Arabidopsis thaliana]

Seq. No. 10867

Contig ID 19857 1.R1040

5'-most EST LIB3170-051-Q1-J1-D10

Seq. No. 10868

Contig ID 19858 1.R1040 5'-most EST sat701010221.h2

Seq. No. 10869

Contig ID 19859 1.R1040 5'-most EST uxk700671202.h1

Method BLASTX
NCBI GI g4337207
BLAST score 786
E value 9.0e-84
Match length 268
% identity 56

NCBI Description (AC006403) putative zinc-finger protein [Arabidopsis

thaliana]

Seq. No. 10870

Contig ID 19868_1.R1040 5'-most EST trc700564095.h1

Method BLASTX
NCBI GI g1871185
BLAST score 300
E value 4.0e-27
Match length 67



```
% identity
                  (U90439) seven in absentia isolog [Arabidopsis thaliana]
NCBI Description
                  10871
Seq. No.
Contig ID
                  19873 1.R1040
                  jC-gmro02910007d04a1
5'-most EST
                  BLASTX
Method
                  g3928084
NCBI GI
BLAST score
                  394
                  3.0e-38
E value
Match length
                  140
                   59
% identity
                   (AC005770) retrotransposon-like protein [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   10872
                   19878 1.R1040
Contig ID
5'-most EST
                  pcp700988751.hl
                   BLASTX
Method
                   g2388580
NCBI GI
BLAST score
                   549
E value
                   1.0e-56
                   112
Match length
% identity
                   (AC000098) Similar to Sequence 10 from patent 5477002
NCBI Description
                   (gb 1253956). [Arabidopsis thaliana]
                   10873
Seq. No.
                   19884 1.R1040
Contig ID
                   q5666694
5'-most EST
Seq. No.
                   10874
                   19888 1.R1040
Contig ID
                   gsv701049390.h1
5'-most EST
                   BLASTX
Method
                   g4204304
NCBI GI
                   252
BLAST score
                   2.0e-21
E value
                   97
Match length
                   53
% identity
                   (AC003027) lcl_prt_seq No definition line found
NCBI Description
                   [Arabidopsis thaliana]
                   10875
Seq. No.
                   19892 1.R1040
Contig ID
                   LIB3050-019-Q1-K1-E6
5'-most EST
                   BLASTX
Method
                   g3461820
NCBI GI
BLAST score
                   190
```

2.0e-14 E value Match length 58 57 % identity

(AC004138) unknown protein [Arabidopsis thaliana] NCBI Description

10876 Seq. No.

Contig ID 19894 1.R1040 5'-most EST $ncj70\overline{0}983051.h1$



10877 Seq. No. 19897 1.R1040 Contig ID 5'-most EST uC-qmflminsoy058b03b1 Seq. No. 10878 19897 2.R1040 Contig ID LIB3093-058-Q1-K1-F6 5'-most EST BLASTX Method g3142290 NCBI GI BLAST score 295 3.0e-26 E value 91 Match length % identity (AC002411) Contains similarity to gb Z69902 from C. NCBI Description elegans. [Arabidopsis thaliana] 10879 Seq. No. 19899 1.R1040 Contig ID 5'-most EST rlr700901811.hl BLASTX Method NCBI GI a3184098 1118 BLAST score 1.0e-122 E value 510 Match length 47 % identity (AL023777) coenzyme a synthetase [Schizosaccharomyces NCBI Description pombe] 10880 Seq. No. 19899 2.R1040 Contig ID LIB3092-008-Q1-K1-F12 5'-most EST Method BLASTX NCBI GI q3184098 BLAST score 368 E value 4.0e-35 Match length 181 % identity (AL023777) coenzyme a synthetase [Schizosaccharomyces NCBI Description pombe] Seq. No. 10881 Contig ID 19905 1.R1040 uC-gmrominsoy047e11b15'-most EST BLASTX Method g4580517 NCBI GI BLAST score 662 E value 2.0e-69 Match length 213 % identity 60 (AF036302) scarecrow-like 5 [Arabidopsis thaliana] NCBI Description Seq. No. 10882 19913 1.R1040 Contig ID $LIB30\overline{5}0-019-Q1-K1-B4$ 5'-most EST

Method BLASTX NCBI GI g2281098



BLAST score 412 3.0e-40E value 125 Match length % identity NCBI Description (AC002333) hypothetical protein [Arabidopsis thaliana] 10883 Seq. No. 19916 1.R1040 Contig ID $rca70\overline{0}998019.h1$ 5'-most EST BLASTX Method g4530126 680

NCBI GI BLAST score 4.0e-71 E value 326 Match length 45 % identity

(AF078082) receptor-like protein kinase homolog RK20-1 NCBI Description

[Phaseolus vulgaris]

Seq. No. 10884

19916 3.R1040 Contig ID $smc70\overline{0}745880.h1$ 5'-most EST

10885 Seq. No.

19918 1.R1040 Contig ID

LIB3050-019-Q1-K1-C1 5'-most EST

Seq. No. 10886

19927 1.R1040 Contig ID

LIB3109-035-Q1-K2-H6 5'-most EST

10887 Seq. No.

Contig ID 19930 1.R1040

LIB3050-019-Q1-K1-E1 5'-most EST

Method BLASTX NCBI GI g3164115 BLAST score 337 E value 3.0e-31 Match length 153 45 % identity

(AJ224145) major latex-like protein [Rubus idaeus] NCBI Description

10888 Seq. No.

19932 1.R1040 Contig ID

LIB3050-019-Q1-K1-B12 5'-most EST

10889 Seq. No.

19932 2.R1040 Contig ID 5'-most EST zsg701117871.h2

10890 Seq. No.

19935 1.R1040 Contig ID

 $LIB30\overline{5}1-035-Q1-K1-B3$ 5'-most EST

Method BLASTX g3924596 NCBI GI 632 BLAST score 8.0e-66 E value 144 Match length



% identity (AF069442) putative phospho-ser/thr phosphatase NCBI Description [Arabidopsis thaliana] 10891 Seq. No. 19938 1.R1040 Contig ID jC-qmf102220069b05d1 5'-most EST BLASTX Method g1706082 NCBI GI BLAST score 144 7.0e-09 E value Match length 46 63 % identity SERINE CARBOXYPEPTIDASE II-3 PRECURSOR (CP-MII.3) NCBI Description 3.4.16.1) II-3 - barley >gi_619350_bbs_153536 CP-MII.3=serine carboxypeptidase [Hordeum vulgare=barley,

>gi_629787_pir__S44191 serine-type carboxypeptidase (EC

cv. Alexis, aleurone, Peptide, 516 aa]

>gi 474392_emb_CAA55478_ (X78877) serine carboxylase II-3

[Hordeum vulgare]

10892 Seq. No. 19939 1.R1040 Contig ID

 $LIB30\overline{5}0-018-Q1-E1-H7$ 5'-most EST

Seq. No. 10893

19943 1.R1040 Contig ID

uC-gmflminsoy023d08b1 5'-most EST

Method BLASTX q3935165 NCBI GI BLAST score 294 2.0e-26 E value 132 Match length % identity 51

NCBI Description (AC004557) F17L21.8 [Arabidopsis thaliana]

Seq. No.

19943 2.R1040 Contig ID

5'-most EST LIB3167-077-P1-K2-H4

10894

BLASTN Method NCBI GI g3228389 BLAST score 38 6.0e-12 E value Match length 74

88 % identity

Genomic sequence for Arabidopsis thaliana BAC F17L21, NCBI Description

complete sequence [Arabidopsis thaliana]

10895 Seq. No.

Contig ID 19958 1.R1040 5'-most EST bth700845494.h1

BLASTX Method a3367536 NCBI GI BLAST score 435 7.0e-43E value Match length 117 % identity 76



NCBI Description (AC004392) Contains similarity to symbiosis-related like

protein F1N20.80 gi_2961343 from A. thaliana BAC gb_AL022140. EST gb_T04695 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 10896

Contig ID 19958_2.R1040 5'-most EST asn701138380.h1

Method BLASTX
NCBI GI g3367536
BLAST score 432
E value 2.0e-42
Match length 117

% identity 75

NCBI Description (AC004392) Contains similarity to symbiosis-related like

protein F1N20.80 gi_2961343 from A. thaliana BAC gb_AL022140. EST gb_T04695 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 10897

Contig ID 19958 3.R1040

5'-most EST LIB3106-023-Q1-K1-B5

Seq. No. 10898

Contig ID 19958_4.R1040

5'-most EST LIB3093-017-Q1-K1-F10

Seq. No. 10899

Contig ID 19963_1.R1040

5'-most EST LIB3050-018-Q1-E1-G12

Seq. No. 10900

Contig ID 19978 1.R1040 5'-most EST uxk700672945.h1

Seq. No. 10901

Contig ID 19986_1.R1040 5'-most EST pxt700943425.h1

Seq. No. 10902

Contig ID 19995_1.R1040 5'-most EST hyd700729446.h1

Method BLASTX
NCBI GI g3033377
BLAST score 1123
E value 1.0e-123
Match length 501
% identity 47

NCBI Description (AC004238) putative berberine bridge enzyme [Arabidopsis

thaliana]

Seq. No. 10903

Contig ID 19995_2.R1040 5'-most EST fua701037134.h1

Method BLASTX
NCBI GI g3033377
BLAST score 314



8.0e-29 E value Match length 131 % identity 46 (AC004238) putative berberine bridge enzyme [Arabidopsis NCBI Description thaliana] 10904 Seq. No. 19997 1.R1040 Contig ID zsq701120155.h1 5'-most EST Method BLASTX g1931647 NCBI GI BLAST score 2216 0.0e + 00E value 564 Match length 74 % identity (U95973) endomembrane protein EMP70 precusor isolog NCBI Description [Arabidopsis thaliana] Seq. No. 10905 19997 2.R1040 Contig ID 5'-most EST LIB3051-087-Q1-K1-A6 Seq. No. 10906 19997 3.R1040 Contig ID LIB3170-059-Q1-K1-A9 5'-most EST Method BLASTX NCBI GI g1931647 BLAST score 231 4.0e-24 E value 90 Match length 59 % identity (U95973) endomembrane protein EMP70 precusor isolog NCBI Description [Arabidopsis thaliana] 10907 Seq. No. 20000 1.R1040 Contig ID LIB3050-018-Q1-E1-C3 5'-most EST Seq. No. 10908 Contig ID 20003 1.R1040 5'-most EST LIB3109-032-Q1-K1-G2 Method BLASTX NCBI GI g3695392 BLAST score 424 4.0e-41E value Match length 201 % identity 47

(AF096371) No definition line found [Arabidopsis thaliana] NCBI Description

10909 Seq. No.

Contig ID 20013 1.R1040

5'-most EST LIB3050-018-Q1-E1-A2

10910 Seq. No.

Contig ID 20016 1.R1040 5'-most EST $crh70\overline{0}852015.h1$

Method BLASTX



NCBI GI g3721540
BLAST score 2338
E value 0.0e+00
Match length 519
% identity 82

NCBI Description (D83583) Sulfite Reductase [Nicotiana tabacum]

>qi 3738234 dbj BAA33796 (AB010717) sulfite reductase

[Nicotiana tabacum] .

Seq. No. 10911

Contig ID 20017 1.R1040

5'-most EST LIB3050-015-Q1-E1-G5

Method BLASTX
NCBI GI g1922242
BLAST score 250
E value 3.0e-21
Match length 122
% identity 54

NCBI Description (Y10084) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10912

Contig ID 20017 2.R1040

5'-most EST LIB3050-018-Q1-E1-A6

Method BLASTX
NCBI GI g1922242
BLAST score 223
E value 5.0e-18
Match length 76
% identity 59

NCBI Description (Y10084) hypothetical protein [Arabidopsis thaliana]

Seq. No. 10913

Contig ID 20017_3.R1040 5'-most EST wrg700787203.h2

Seq. No. 10914

Contig ID 20017_4.R1040

5'-most EST LIB3055-012-Q1-N1-G8

Seq. No. 10915

Contig ID 20021_1.R1040

5'-most EST LIB3050-018-Q1-E1-B1

Method BLASTX
NCBI GI g4063750
BLAST score 140
E value 1.0e-08
Match length 90
% identity 42

NCBI Description (AC005851) putative indole-3-acetate

beta-glucosyltransferase synthetase [Arabidopsis thaliana]

Seq. No. 10916

Contig ID 20024 1.R1040

5'-most EST LIB3050-018-Q1-E1-B12

Seq. No. 10917

Contig ID 20025_1.R1040

10922

```
5'-most EST
                   uC-gmropic042c12b1
Method
                   BLASTX
                   g2842485
NCBI GI
BLAST score
                   211
                   3.0e-16
E value
Match length
                   118
% identity
                   38
NCBI Description
                  (AL021749) hypothetical protein [Arabidopsis thaliana]
                   10918
Seq. No.
Contig ID
                   20030 1.R1040
5'-most EST
                   LIB3050-017-Q1-E1-G2
Method
                   BLASTX
NCBI GI
                   g282963
BLAST score
                   629
E value
                   1.0e-65
Match length
                   176
% identity
                   70
NCBI Description
                  transforming protein (myb) homolog (clone myb.Ph2) - garden
                   petunia >gi 20561 emb CAA78387 (Z13997) protein 2 [Petunia
                   x hybrida]
Seq. No.
                   10919
Contig ID
                   20038 1.R1040
5'-most EST
                   uC-gmrominsoy282f05b1
Method
                   BLASTX
NCBI GI
                   g2275214
BLAST score
                   232
E value
                   3.0e-19
Match length
                   91
% identity
                   59
NCBI Description
                  (AC002337) hypothetical protein [Arabidopsis thaliana]
                   10920
Seq. No.
Contig ID
                   20040 1.R1040
5'-most EST
                   vzy700754715.h1
Method
                  BLASTX
NCBI GI
                   g3646451
BLAST score
                   204
E value
                   8.0e-16
Match length
                   112
% identity
                   45
                   (AL031603) mRNA cap methyltransferase [Schizosaccharomyces
NCBI Description
                  pombe]
Seq. No.
                   10921
Contig ID
                   20043 1.R1040
5'-most EST
                   zsg70\overline{1}124088.h1
Method
                  BLASTX
NCBI GI
                  q2828147
BLAST score
                   537
E value
                  2.0e-54
Match length
                  222
% identity
                  50
NCBI Description (AF042384) BC-2 protein [Homo sapiens]
```



```
Contig ID
                   20043 2.R1040
5'-most EST
                   LIB3139-097-P1-N1-A11
Method
                   BLASTX
NCBI GI
                   g3021268
BLAST score
                   114
E value
                   3.0e-11
Match length
                   193
% identity
                   36
NCBI Description (AL022347) putative protein [Arabidopsis thaliana]
Seq. No.
                   10923
Contig ID
                   20043 7.R1040
5'-most EST
                   eep700865965.hl
Seq. No.
                   10924
                   20047 1.R1040
Contig ID
5'-most EST
                   g4313881
Seq. No.
                   10925
Contig ID
                   20051 1.R1040
5'-most EST
                   trc700562631.h1
Method
                   BLASTX
NCBI GI
                   q1732247
BLAST score
                   329
                   3.0e-37
E value
Match length
                   121
% identity
                   64
NCBI Description (U72762) transcription factor Myb1 [Nicotiana tabacum]
Seq. No.
                  10926
Contig ID
                   20051 2.R1040
5'-most EST
                  trc700566792.h1
Method
                  BLASTX
NCBI GI
                   q1945279
BLAST score
                   368
E value
                  2.0e-35
Match length
                  96
% identity
                   64
NCBI Description (Y11350) myb factor [Oryza sativa]
                  10927
Seq. No.
                  20051 3.R1040
Contig ID
5'-most EST
                  uC-gmflminsoy043b11b1
Method
                  BLASTX
NCBI GI
                  g1946267
BLAST score
                  247
E value
                  2.0e-21
Match length
                  61
% identity
                  70
NCBI Description (Y11415) myb [Oryza sativa]
```

Contig ID 20056_1.R1040 5'-most EST rlr700898958.h1

Method BLASTX NCBI GI g1168940 BLAST score 480

```
E value
                  5.0e-48
Match length
                  175
% identity
                  51
NCBI Description
                  CHORISMATE MUTASE PRECURSOR (CM-1) >gi 629509 pir S38958
                  chorismate mutase precursor - Arabidopsis thaliana
                  >gi 429153 emb CAA81286 (Z26519) chorismate mutase
                  precursor [Arabidopsis thaliana]
                  10929
Seq. No.
Contig ID
                  20056 2.R1040
5'-most EST
                  gsv701047857.h1
Method
                  BLASTX
                  g1168940
```

NCBI GI g1168940
BLAST score 157
E value 2.0e-10
Match length 96
% identity 39

NCBI Description CHORISMATE MUTASE PRECURSOR (CM-1) >gi 629509 pir S38958

chorismate mutase precursor - Arabidopsis thaliana >gi_429153 emb_CAA81286 (Z26519) chorismate mutase

precursor [Arabidopsis thaliana]

Seq. No. 10930

Contig ID 20063_1.R1040

5'-most EST LIB30 $\overline{5}$ 0-017-Q1-E1-F6

Method BLASTX
NCBI GI g1730502
BLAST score 271
E value 1.0e-23
Match length 116
% identity 47

NCBI Description TRANSMEMBRANE PROTEIN PFT27 >gi 110903 pir A31351 probable

transmembrane protein FT27 - mouse >qi 535682 (M23568)

transmembrane protein [Mus musculus]

Seq. No. 10931

Contig ID 20065 1.R1040

5'-most EST LIB3050-017-Q1-E1-F9

Method BLASTX
NCBI GI g1362019
BLAST score 166
E value 2.0e-11
Match length 68
% identity 56

NCBI Description zinc finger protein 6 - Arabidopsis thaliana >qi 790683

(L39649) zinc finger protein [Arabidopsis thaliana]

Seq. No. 10932

Contig ID 20067_1.R1040 5'-most EST wrg700787082.h2

Method BLASTX
NCBI GI g3757527
BLAST score 472
E value 4.0e-47
Match length 116
% identity 73

NCBI Description (AC005167) putative ribosomal protein L27 [Arabidopsis

% identity



thaliana]

```
10933
Seq. No.
Contig ID
                  20067 3.R1040
5'-most EST
                  wrg700787493.h2
                  10934
Seq. No.
                   20068 1.R1040
Contig ID
                  LIB3050-012-Q1-E1-C3
5'-most EST
                   10935
Seq. No.
                   20070 1.R1040
Contig ID
                   uC\text{-}gm\bar{r}ominsoy309h05b1
5'-most EST
Method
                   BLASTX
                   g2642158
NCBI GI
BLAST score
                   739
                   3.0e-78
E value
Match length
                   236
                   60
% identity
                   (AC003000) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   10936
                   20072 1.R1040
Contig ID
                   LIB3109-016-Q1-K1-F10
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3128173
BLAST score
                   415
E value
                   7.0e-41
Match length
                   97
                   79
% identity
                  (AC004521) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   10937
Seq. No.
                   20081 1.R1040
Contiq ID
                   LIB3139-006-P1-N1-A11
5'-most EST
                   BLASTX
Method
                   g4220476
NCBI GI
                   310
BLAST score
                   5.0e-48
E value
Match length
                   140
% identity
                   62
                   (AC006069) ribophorin I-like protein [Arabidopsis thaliana]
NCBI Description
                   10938
Seq. No.
                   20085 1.R1040
Contig ID
                   jC-gmst02400056c10d1
5'-most EST
                   10939
Seq. No.
                   20085 2.R1040
Contig ID
                   zsq701127688.h1
5'-most EST
Method
                   BLASTX
                   q2832658
NCBI GI
                   300
BLAST score
                   1.0e-26
E value
Match length
                   120
```

NCBI Description (AL021710) putative protein [Arabidopsis thaliana]



20085 3.R1040 Contig ID hrw701062022.h1 5'-most EST

10941 Seq. No.

Contig ID 20085 4.R1040 hyd700727277.h1 5'-most EST

BLASTX Method q4432845 NCBI GI BLAST score 195 3.0e-15 E value 98 Match length 43 % identity

NCBI Description (AC006283) unknown protein [Arabidopsis thaliana]

Seq. No. 10942

20085 6.R1040 Contig ID 5'-most EST zsg701129657.h1

Seq. No. 10943

20086 1.R1040 Contig ID

5'-most EST LIB3050-017-Q1-E1-A10

Seq. No. 10944

Contig ID 20090 2.R1040 5'-most EST awf700836953.hl

Seq. No. 10945

20091 1.R1040 Contig ID

5'-most EST g5126401 Method BLASTX NCBI GI g2160182 273 BLAST score 1.0e-23 E value 148 Match length

46 % identity

NCBI Description (AC000132) ESTs gb ATTS1236, gb T43334, gb N97019, gb AA395203

come from this gene. [Arabidopsis thaliana]

Seq. No. 10946

Contig ID 20091 2.R1040

5'-most EST LIB3051-006-Q1-K1-D4

Method BLASTX NCBI GI q2160182 BLAST score 156 2.0e-10 E value Match length 114 % identity

NCBI Description (AC000132) ESTs gb_ATTS1236, gb_T43334, gb_N97019, gb_AA395203

come from this gene. [Arabidopsis thaliana]

Seq. No. 10947

Contig ID 20096 1.R1040

5'-most EST LIB3055-011-Q1-N1-H7

Method BLASTX NCBI GI g4056438



BLAST score 248 E value 8.0e-21 Match length 144 % identity 41

NCBI Description (AC005990) F508.11 [Arabidopsis thaliana]

Seq. No. 109

Contig ID 20096_2.R1040 5'-most EST uxk700670789.h1

Seq. No. 10949

Contig ID 20097_1.R1040

5'-most EST LIB3050-017-Q1-E1-B12

Seq. No. 10950

Contig ID 20097_2.R1040 5'-most EST bth700846843.h1

Seq. No. 10951

Contig ID 20105_1.R1040

5'-most EST LIB3050-016-Q1-E1-G9

Seq. No. 10952

Contig ID 20106 1.R1040

5'-most EST LIB3050-016-Q1-E1-G1

Seq. No. 10953

Contig ID 20109_1.R1040 5'-most EST pxt700943385.h1

Seq. No. 10954

Contig ID 20115_1.R1040

5'-most EST LIB3109-042-Q1-K1-A1

Method BLASTX
NCBI GI g1351365
BLAST score 153
E value 5.0e-10
Match length 48
% identity 60

NCBI Description UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 6.7 KD PROTEIN

(CR6) >gi_2130002_pir__S68969 ubiquinol--cytochrome-c reductase (EC 1.10.2.2) - potato >gi_633683_emb_CAA57768_ (X82325) cytochrome c reductase subunit [Solanum tuberosum]

Seq. No. 10955

Contig ID 20122 1.R1040

NCBI Description (AL021811) putative protein [Arabidopsis thaliana]

Seq. No. 10956

Contig ID 20127 1.R1040



```
g5677300
5'-most EST
                  BLASTX
Method
                  g4454452
NCBI GI
                  406
BLAST score
E value
                  2.0e-39
Match length
                  91
% identity
                  (AC006234) unknown protein [Arabidopsis thaliana]
NCBI Description
                  10957
Seq. No.
                   20127 2.R1040
Contig ID
                  ASG3244V4L-01-Q1-E1-H2
5'-most EST
                   10958
Seq. No.
                   20145 1.R1040
Contig ID
5'-most EST
                   vwf700676069.h1
Method
                   BLASTX
NCBI GI
                   q3702332
                   179
BLAST score
                   8.0e-13
E value
                   98
Match length
% identity
NCBI Description (AC005397) unknown protein [Arabidopsis thaliana]
                   10959
Seq. No.
                   20147 1.R1040
Contig ID
5'-most EST
                   zsq701126402.h1
                   BLASTX
Method
                   q3850108
NCBI GI
BLAST score
                   437
                   3.0e-43
E value
                   147
Match length
% identity
                   56
                   (AL033388) putative calcium-transporting atpase
NCBI Description
                   [Schizosaccharomyces pombe]
                   10960
Seq. No.
                   20155 1.R1040
Contig ID
                   LIB3050-016-Q1-E1-B1
 5'-most EST
Method
                   BLASTX
 NCBI GI
                   q1853968
                   568
 BLAST score
                   2.0e-58
 E value
                   154
 Match length
                   71
 % identity
                   (D88121) CPRD12 protein [Vigna unguiculata]
 NCBI Description
                   10961
 Seq. No.
                   20163 1.R1040
 Contig ID
                   LIB3093-031-Q1-K1-F2
 5'-most EST
                   10962
 Seq. No.
                   20167 1.R1040
 Contig ID
                   jC-gmro02910054f07d1
 5'-most EST
```

1904

10963

20168 1.R1040

Seq. No.

Contig ID



```
LIB3109-054-Q1-K1-F4
5'-most EST
                  BLASTX
Method
                  q1730502
NCBI GI
                  163
BLAST score
                  8.0e-11
E value
Match length
                  81
% identity
                  TRANSMEMBRANE PROTEIN PFT27 >gi_110903_pir__A31351 probable
NCBI Description
                   transmembrane protein FT27 - mouse >gi_535682 (M23568)
                   transmembrane protein [Mus musculus]
                   10964
Seq. No.
                   20168 2.R1040
Contig ID
                   jC-gmle01810082f07a1
5'-most EST
                   10965
Seq. No.
Contig ID
                   20180 1.R1040
5'-most EST
                   dpv701098046.h1
                   10966
Seq. No.
                   20182 1.R1040
Contig ID
                   uC-gmrominsoy293b01b1
5'-most EST
                   BLASTX
Method
                   g4558678
NCBI GI
                   256
BLAST score
                   2.0e-21
E value
                   119
Match length
                   49
% identity
NCBI Description (AC006586) unknown protein [Arabidopsis thaliana]
                   10967
Seq. No.
                   20183 1.R1040
Contig ID
                   LIB30\overline{5}0-016-Q1-E1-A3
5'-most EST
                   10968
Seq. No.
                   20184 1.R1040
Contig ID
5'-most EST
                   LIB3050-010-Q1-E1-H4
                   10969
Seq. No.
                   20185 1.R1040
Contig ID
5'-most EST
                   LIB3050-016-Q1-E1-A5
Seq. No.
                   10970
                   20188 1.R1040
Contig ID
                   uC-qmflminsoy021e02b1
 5'-most EST
                   BLASTX
Method
                   q2369766
NCBI GI
BLAST score
                   1153
                    1.0e-126
E value
                    303
Match length
 % identity
                   (AJ001304) hypothetical protein [Citrus x paradisi]
 NCBI Description
```

Contig ID 20188_2.R1040

5'-most EST LIB3106-026-Q1-K1-E1

Method BLASTX

```
g2369766
NCBI GI
BLAST score
                   217
                   2.0e-17
E value
Match length
                   130
% identity
                   (AJ001304) hypothetical protein [Citrus x paradisi]
NCBI Description
                   10972
Seq. No.
                   20188 5.R1040
Contig ID
5'-most EST
                   kl1701210213.h1
                   10973
Seq. No.
                   20194 1.R1040
Contig ID
                   LIB3050-015-Q1-E1-F7
5'-most EST
                   BLASTX
Method
NCBI GI
                   q642134
BLAST score
                   431
E value
                   2.0e-42
                   138
Match length
                   62
% identity
                   (D45355) protein kinase [Arabidopsis thaliana]
NCBI Description
                   >qi 3063704 emb CAA18595.1 (AL022537) protein kinase AME3
                   [Arabidopsis thaliana]
                   10974
Seq. No.
                   20194 4.R1040
Contig ID
5'-most EST
                   gsv701054279.h1
                   10975
Seq. No.
Contig ID
                   20202 1.R1040
                   fua701038834.h1
5'-most EST
Method
                   BLASTX
                   q452593
NCBI GI
BLAST score
                   1056
                   1.0e-115
E value
Match length
                   291
% identity
                   67
                   (D21814) ORF [Lilium longiflorum]
NCBI Description
                   10976
Seq. No.
                   20209 1.R1040
Contig ID
                   g4295899
5'-most EST
                   10977
Seq. No.
                   20211 1.R1040
Contig ID
                   smc70\overline{0}746358.h1
5'-most EST
                   BLASTX
Method
                   g3550519
NCBI GI
                    1317
BLAST score
                    1.0e-151
E value
                    335
Match length
                   78
 % identity
```

NCBI Description (AJ007630) oxygenase [Nicotiana tabacum]

Seq. No. 10978

Contig ID 20213_1.R1040 5'-most EST vzy700756512.h1



```
Method BLASTX
NCBI GI g3335335
BLAST score 472
E value 3.0e-47
Match length 121
% identity 73
```

NCBI Description (AC004512) ESTs gb_F14113 and gb_T42122 come from this region. [Arabidopsis thaliana]

region. [Alabidopsis ch

Seq. No. 10979

Contig ID 20215_1.R1040

5'-most EST uC-gmrominsoy208b07b1

Method BLASTX
NCBI GI g121345
BLAST score 1683
E value 0.0e+00
Match length 341
% identity 91

NCBI Description GLUTAMINE SYNTHETASE PR-2 (ISOZYME ALPHA)

(GLUTAMATE--AMMONIA LIGASE) >gi_68593_pir_AJFBQA

glutamate--ammonia ligase (EC 6.3.1.2) alpha, cytosolic - kidney bean >gi 21013 emb CAA27632 (X04002) glutamine synthetase subunit (aa 1-356) [Phaseolus vulgaris] >gi 225068 prf 1208270B synthetase R2,Gln [Phaseolus

vulgaris]

Seq. No. 10980

Contig ID 20215_2.R1040

5'-most EST LIB3087-013-Q1-K1-C5

Method BLASTN
NCBI GI g21012
BLAST score 366
E value 0.0e+00
Match length 762
% identity 93

NCBI Description Phaseolus vulgaris mRNA (pR-2) for glutamine synthetase

(GS, EC 6.3.1.2)

Seq. No. 10981

Contig ID 20217 1.R1040

5'-most EST LIB $30\overline{5}0-015-Q1-E1-E12$

Method BLASTX
NCBI GI g1491776
BLAST score 336
E value 1.0e-31
Match length 87
% identity 78

NCBI Description (M37636) cationic peroxidase [Arachis hypogaea]

Seq. No. 10982

Contig ID 20228_1.R1040 5'-most EST rca701001008.h1

Method BLASTX
NCBI GI g4567226
BLAST score 195
E value 1.0e-14
Match length 69



% identity 61

NCBI Description (AC007119) unknown protein [Arabidopsis thaliana]

Seq. No. 10983

Contig ID 20230_1.R1040 5'-most EST zhf700954844.h1

Method BLASTX
NCBI GI g3688174
BLAST score 537
E value 8.0e-55
Match length 148
% identity 68

NCBI Description (AL031804) putative protein [Arabidopsis thaliana]

Seq. No. 10984

Contig ID 20239_1.R1040 5'-most EST pxt700946086.h1

Method BLASTX
NCBI GI g3064039
BLAST score 323
E value 1.0e-29
Match length 142
% identity 41

NCBI Description (AF054445) major latex protein homolog [Mesembryanthemum

crystallinum]

Seq. No. 10985

Contig ID 20242 1.R1040

5'-most EST uC-gmrominsoy155d05b1

Seq. No. 10986

Contig ID 20243 1.R1040

5'-most EST LIB $30\overline{5}1-048-Q1-K1-E9$

Seq. No. 10987

Contig ID 20249_1.R1040 5'-most EST vwf700676351.h1

Method BLASTX
NCBI GI g3860250
BLAST score 400
E value 7.0e-39
Match length 163
% identity 53

NCBI Description (AC005824) putative chloroplast prephenate dehydratase

[Arabidopsis thaliana]

Seq. No. 10988

Contig ID 20254 1.R1040 5'-most EST zzp700833517.h1

Method BLASTX
NCBI GI g4103635
BLAST score 432
E value 2.0e-42
Match length 162
% identity 56

NCBI Description (AF026538) ABA-responsive protein [Hordeum vulgare]



```
Seq. No.
                   10989
Contig ID
                   20262 1.R1040
5'-most EST
                  LIB3050-015-Q1-E1-A1
Method
                  BLASTX
NCBI GI
                  g2460188
BLAST score
                  1082
E value
                  1.0e-118
Match length
                  320
% identity
                  66
                  (AF020785) polygalacturonase inhibiting protein [Prunus
NCBI Description
                  armeniaca]
Seq. No.
                  10990
                  20268 1.R1040
Contig ID
                  LIB3050-014-Q1-E1-G6
5'-most EST
Method
                  BLASTX
NCBI GI
                  q3618214
BLAST score
                  203
                  2.0e-15
E value
                  186
Match length
% identity
NCBI Description
                  (AL031579) dihydrofolate reductase [Schizosaccharomyces
                  pombe]
                  10991
Seq. No.
Contig ID
                  20273 1.R1040
5'-most EST
                  asn701142539.h1
                  10992
Seq. No.
Contig ID
                  20275 1.R1040
5'-most EST
                  fC-qmse700839629a5
Method
                  BLASTX
                  g3152608
NCBI GI
BLAST score
                   488
                  3.0e-49
E value
Match length
                  96
% identity
                   94
                  (AC004482) putative BEL1-like homeotic protein [Arabidopsis
NCBI Description
                  thalianal
Seq. No.
                   10993
                  20278 1.R1040
Contig ID
5'-most EST
                  sat701010383.h1
Method
                  BLASTX
NCBI GI
                  g3193287
BLAST score
                  770
E value
                  3.0e-82
Match length
                  156
                  88
% identity
NCBI Description
                  (AF069298) Arabidopsis predicted protein of unknown
                  function T10P11.19 (GB:AC002330) [Arabidopsis thaliana]
```

Seq. No. 10994 Contig ID 20281 1.R1040

5'-most EST jC-gmle01810035g12a2

Seq. No. 10995



Contig ID 20281_2.R1040 5'-most EST seb700653544.h1

Seq. No. 10996

Contig ID 20288_1.R1040

5'-most EST LIB3170-055-Q1-K1-A3

Method BLASTX
NCBI GI 94469009
BLAST score 491
E value 3.0e-49
Match length 137
% identity 75

NCBI Description (AL035602) putative protein [Arabidopsis thaliana]

Seq. No. 10997

Contig ID 20293_1.R1040 5'-most EST fde700871956.h1

Seq. No. 10998

Contig ID 20296 1.R1040

5'-most EST LIB30 $\overline{5}$ 0-014-Q1-E1-F8

Method BLASTN
NCBI GI g1015315
BLAST score 66
E value 9.0e-29
Match length 82

% identity 95

NCBI Description Pisum sativum (clone PsRCI35-2) ribosomal protein L41 mRNA,

complete cds

Seq. No. 10999

Contig ID 20297_1.R1040

5'-most EST LIB3050-014-Q1-E1-F9

Seq. No. 11000

Contig ID 20300_1.R1040 5'-most EST uC-gmropic030b10b1

Seq. No. 11001

Contig ID 20300 2.R1040

5'-most EST LIB3051-018-Q1-E1-H2

Seq. No. 11002

Contig ID 20305_1.R1040 5'-most EST uC-gmropic102d02b1

Method BLASTX
NCBI GI g4538939
BLAST score 1181
E value 1.0e-130
Match length 254
% identity 87

NCBI Description (AL049483) Col-O casein kinase I-like protein [Arabidopsis

thaliana]

Seq. No. 11003

Contig ID 20308 1.R1040

5'-most EST LIB3050-014-Q1-E1-D12



```
BLASTX
Method
                  q2642215
NCBI GI
                  304
BLAST score
                  2.0e-27
E value
Match length
                  79
% identity
                  70
NCBI Description (AF030386) NOI protein [Arabidopsis thaliana]
                  11004
Seq. No.
                  20312 1.R1040
Contig ID
5'-most EST
                  LIB3092-016-Q1-K1-G8
                   11005
Seq. No.
                   20313 1.R1040
Contig ID
                  LIB3050-003-Q1-E1-F3
5'-most EST
Seq. No.
                   11006
                   20313 2.R1040
Contig ID
                   LIB3050-014-Q1-E1-D7
5'-most EST
                   11007
Seq. No.
                   20316 1.R1040
Contig ID
                   LIB3050-014-Q1-E1-E11
5'-most EST
                   11008
Seq. No.
                   20317 1.R1040
Contig ID
5'-most EST
                   LIB3050-014-Q1-E1-C3
                   BLASTX
Method
                   g3335376
NCBI GI
BLAST score
                   300
                   2.0e-27
E value
                   73
Match length
% identity
                   (AC003028) putative ammonium transporter [Arabidopsis
NCBI Description
                   thaliana]
                   11009
Seq. No.
                   20319 1.R1040
Contig ID
                   LIB3093-009-Q1-K1-E5
 5'-most EST
Method
                   BLASTX
NCBI GI
                   g3281848
 BLAST score
                   467
 E value
                   4.0e-46
                   251
 Match length
                   49
 % identity
                   (AL031004) putative protein [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
                   11010
                   20319 2.R1040
 Contig ID
                   LIB3050-014-Q1-E1-A6
 5'-most EST
                   11011
 Seq. No.
```

Contig ID 20319_3.R1040

5'-most EST LIB3053-014-Q1-N1-H2

Method BLASTX
NCBI GI g4559352
BLAST score 148

1911



```
3.0e-09
E value
Match length
                  83
% identity
                  (AC006585) putative DNA binding protein [Arabidopsis
NCBI Description
                  thaliana]
                  11012
Seq. No.
                  20319 5.R1040
Contig ID
                  smc700749977.h1
5'-most EST
Seq. No.
                  11013
                  20319 6.R1040
Contig ID
                  q5677508
                  BLASTX
```

5'-most EST Method q3281848 NCBI GI BLAST score 331 E value 2.0e-30 Match length 182 % identity

NCBI Description (AL031004) putative protein [Arabidopsis thaliana]

11014 Seq. No. 20319 7.R1040 Contig ID LIB3050-007-Q1-E1-E1 5'-most EST

11015 Seq. No. 20319 8.R1040 Contig ID 5'-most EST rca701002460.hl

Seq. No. 11016 20319 9.R1040 Contig ID 5'-most EST LIB3139-054-P1-N1-A7 BLASTX

Method g2130024 NCBI GI 216 BLAST score 7.0e-30 E value 151 Match length 49 % identity

DNA-binding protein ABF2 - wild oat NCBI Description

>gi 1159879 emb CAA88331_ (Z48431) DNA-binding protein

[Avena fatua]

11017 Seq. No. 20319 11.R1040 Contig ID pmv700891060.hl 5'-most EST

11018 Seq. No. 20330 1.R1040 Contig ID 5'-most EST fjg700968365.hl

11019 Seq. No. 20330 2.R1040 Contig ID $jsh70\overline{1}070124.h1$ 5'-most EST

Seq. No. 11020 Contig ID 20333 1.R1040

uC-gmflminsoy046e07b1 5'-most EST



Method BLASTX
NCBI GI g3152559
BLAST score 591
E value 8.0e-61
Match length 165
% identity 66

NCBI Description (AC002986) Similarity to A. thaliana gene product

F21M12.20, gb_AC000132. EST gb_Z25651 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 11021

Contig ID 20333_2.R1040 5'-most EST sat701011980.h1

Seq. No. 11022

Contig ID 20334_1.R1040

5'-most EST LIB3050-014-Q1-E1-C12

Method BLASTX
NCBI GI g2270994
BLAST score 159
E value 8.0e-11
Match length 57
% identity 53

NCBI Description (AF004809) Ca+2-binding EF hand protein [Glycine max]

Seq. No. 11023

Contig ID 20340_1.R1040

5'-most EST LIB3050-012-Q1-E1-E5

Seq. No. 11024

Contig ID 20344_1.R1040 5'-most EST gsv701045374.h1

Method BLASTX
NCBI GI g3420056
BLAST score 220
E value 9.0e-18
Match length 85
% identity 53

NCBI Description (AC004680) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11025

Contig ID 20346_1.R1040

5'-most EST LIB3050-013-Q1-E1-H6

Seq. No. 11026

Contig ID 20348_1.R1040 5'-most EST dpv701098552.h1

Seq. No. 11027

Contig ID 20351_1.R1040

5'-most EST LIB3050-014-Q1-E1-A11

Method BLASTX
NCBI GI g3885339
BLAST score 214
E value 3.0e-17
Match length 128
% identity 45

1913



NCBI Description (AC005623) putative bzip protein [Arabidopsis thaliana]

Seq. No. 11028

Contig ID 20360 1.R1040

5'-most EST LIB3139-075-P1-N1-C2

Seq. No. 11029

Contig ID 20360_2.R1040 5'-most EST zhf700963712.h1

Seq. No. 11030

Contig ID 20361_1.R1040

5'-most EST jC-gmst02400069e10a1

Method BLASTX
NCBI GI g4454033
BLAST score 487
E value 6.0e-49
Match length 213
% identity 47

NCBI Description (AL035394) putative potassium transport protein

[Arabidopsis thaliana]

Seq. No. 11031

Contig ID 20362 1.R1040

5'-most EST LIB3050-013-Q1-E1-F10

Seq. No. 11032

Contig ID 20365_1.R1040

5'-most EST LIB3050-013-Q1-E1-F2

Seq. No. 11033

Contig ID 20368_1.R1040

5'-most EST uC-gmrominsoy298f07b1

Seq. No. 11034

Contig ID 20370 1.R1040 5'-most EST dpv701097078.h1

Seq. No. 11035

Contig ID 20371_1.R1040

5'-most EST jC-gmle01810093c07d1

Method BLASTX
NCBI GI g3426048
BLAST score 239
E value 4.0e-20
Match length 63
% identity 78

NCBI Description (AC005168) putative hydroxymethylglutaryl-CoA lyase

precursor [Arabidopsis thaliana]

Seq. No. 11036

Contig ID 20373_1.R1040

5'-most EST LIB31 $\overline{3}$ 9-069-P1-N1-F4

Seq. No. 11037

Contig ID 20374 1.R1040

5'-most EST LIB3138-029-Q1-N1-F7

% identity

50



```
BLASTX
Method
                  g1922242
NCBI GI
BLAST score
                  200
                  5.0e-15
E value
Match length
                  68
% identity
                  56
                  (Y10084) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  11038
Seq. No.
Contig ID
                  20377 1.R1040
                  LIB3050-013-Q1-E1-C7
5'-most EST
                  BLASTX
Method
                  q2500572
NCBI GI
BLAST score
                  176
                   1.0e-12
E value
Match length
                   139
% identity
                   37
                   RIBONUCLEASE S-2 PRECURSOR (STYLAR GLYCOPROTEIN 2)
NCBI Description
                   (S2-RNASE) >gi_1405424_emb_CAA65319_ (X96465) S2-RNase
                   [Antirrhinum hispanicum]
Seq. No.
                   11039
                   20385 1.R1040
Contig ID
                   sat701002746.h2
5'-most EST
                   BLASTX
Method
                   q4467156
NCBI GI
BLAST score
                   730
                   3.0e-77
E value
Match length
                   218
% identity
                   69
                  (AL035540) putative protein [Arabidopsis thaliana]
NCBI Description
                   11040
Seq. No.
                   20387 1.R1040
Contig ID
                   LIB3139-076-P1-N1-E11
5'-most EST
Seq. No.
                   11041
                   20388 1.R1040
Contig ID
                   LIB3051-048-Q1-K1-H11
5'-most EST
                   BLASTX
Method
                   q2832629
NCBI GI
BLAST score
                   838
                   9.0e-90
E value
                   314
Match length
                   49
% identity
                   (ALO21711) 4-coumarate-CoA ligase - like [Arabidopsis
NCBI Description
                   thaliana]
                   11042
Seq. No.
                   20391 1.R1040
Contig ID
5'-most EST
                   LIB3050-013-Q1-E1-E1
Method
                   BLASTX
                   g1363482
NCBI GI
BLAST score
                   211
                   9.0e-17
E value
                   101
Match length
```

1915



NCBI Description IAA11 protein - Arabidopsis thaliana >gi_972925 (U18413) IAA11 [Arabidopsis thaliana] 11043 Seq. No. 20398 1.R1040 Contig ID $LIB30\overline{5}0-013-Q1-E1-B12$ 5'-most EST BLASTX Method q1706282 NCBI GI 314 BLAST score 6.0e-29 E value 112 Match length 54 % identity DISEASE RESISTANCE RESPONSE PROTEIN 206 >gi 508844 (U11716) NCBI Description disease resistance response protein 206-d [Pisum sativum] 11044 Seq. No. 20401 1.R1040 Contig ID qsv701049038.hl 5'-most EST 11045 Seq. No. 20410 1.R1040 Contig ID trc700565029.hl 5'-most EST 11046 Seq. No. 20410 2.R1040 Contig ID LIB3074-033-Q1-K1-G12 5'-most EST Seq. No. 11047 20412 1.R1040 Contig ID uC-gmrominsoy106e08b1 5'-most EST BLASTX Method q481002 NCBI GI 1406 BLAST score 1.0e-156 E value 338 Match length 78 % identity NCBI Description protein kinase - common ice plant >gi 407057 emb CAA81443 (Z26846) protein kinase [Mesembryanthemum crystallinum] 11048 Seq. No. 20412 2.R1040 Contig ID fC-gmse700842846d3 5'-most EST

Seq. No. 11049

Contig ID 20412 3.R1040

5'-most EST LIB3051-051-Q1-K1-D2

Method BLASTX
NCBI GI g2055374
BLAST score 381
E value 6.0e-37
Match length 87
% identity 85

NCBI Description (U29095) serine-threonine protein kinase [Triticum

aestivum]

Seq. No. 11050

Contig ID 20414_1.R1040

5'-most EST



```
LIB3139-101-P1-N1-D5
5'-most EST
                   11051
Seq. No.
Contig ID
                   20421 1.R1040
5'-most EST
                   LIB3051-020-Q1-E1-A11
Method
                   BLASTX
                   q3738285
NCBI GI
BLAST score
                   482
                   4.0e-48
E value
                   159
Match length
% identity
NCBI Description (AC005309) unknown protein [Arabidopsis thaliana]
                   11052
Seq. No.
                   20421 2.R1040
Contig ID
5'-most EST
                   q4314\overline{0}28
Method
                   BLASTX
NCBI GI
                   q3738285
                   313
BLAST score
                   1.0e-28
E value
                   102
Match length
                   57
% identity
NCBI Description (AC005309) unknown protein [Arabidopsis thaliana]
                   11053
Seq. No.
                   20421 3.R1040
Contig ID
                   rlr700900790.hl
5'-most EST
                   BLASTX
Method
                   q3738285
NCBI GI
BLAST score
                   181
                   3.0e-13
E value
                   57
Match length
% identity
                   (AC005309) unknown protein [Arabidopsis thaliana]
NCBI Description
                   11054
Seq. No.
                   20421 4.R1040
Contig ID
                   LIB3093-030-Q1-K1-H10
5'-most EST
Method
                   BLASTX
                   q3738285
NCBI GI
BLAST score
                   173
                   3.0e-12
E value
Match length
                   87
                   43
% identity
                   (AC005309) unknown protein [Arabidopsis thaliana]
NCBI Description
                   11055
Seq. No.
                   20427_1.R1040
Contig ID
                   LIB3106-010-Q1-K1-D12
5'-most EST
                   11056
Seq. No.
Contig ID
                   20431 1.R1040
                   LIB30\overline{5}0-012-Q1-E1-F4
5'-most EST
                   11057
Seq. No.
                   20433 1.R1040
Contig ID
                   uC-gmrominsoy185f05b1
```



11058 Seq. No. 20435 1.R1040 Contig ID gsv701045964.hl 5'-most EST Method BLASTX NCBI GI q2708747 BLAST score 154 6.0e-10 E value 74 Match length 12 % identity (AC003952) putative glycine-rich, zinc-finger DNA-binding NCBI Description protein [Arabidopsis thaliana] 11059 Seq. No. 20436 1.R1040 Contig ID 5'-most EST LIB3050-012-Q1-E1-G1 11060 Seq. No. 20440 1.R1040 Contig ID LIB3050-012-Q1-E1-G2 5'-most EST BLASTX Method q3885334 NCBI GI 281 BLAST score 1.0e-35 E value 105 Match length % identity (AC005623) putative argonaute protein [Arabidopsis NCBI Description thaliana] Seq. No. 11061 20441 1.R1040 Contiq ID jC-qmst02400065g01d1 5'-most EST Method BLASTX q3309583 NCBI GI BLAST score 540 3.0e-55 E value Match length 112 93 % identity (AF073830) fructose-6-phosphate NCBI Description 2-kinase/fructose-2,6-bisphosphatase [Solanum tuberosum] 11062 Seq. No. 20445 1.R1040 Contig ID LIB3109-007-Q1-K1-G4 5'-most EST BLASTX Method g2760317 NCBI GI 258 BLAST score 5.0e-27 E value 106 Match length % identity 56 (AC002130) F1N21.1 [Arabidopsis thaliana]

NCBI Description

Seq. No. 20445 2.R1040 Contig ID leu701155845.h1 5'-most EST

11063

BLASTX Method g3341697 NCBI GI



BLAST score 237
E value 1.0e-19
Match length 58
% identity 69

NCBI Description (AC003672) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11064

Contig ID 20445 3.R1040

5'-most EST LIB3170-052-Q1-K1-F10

Seq. No. 11065

Contig ID 20445_4.R1040 5'-most EST bth700847369.h1

Seq. No. 11066

Contig ID 20446 1.R1040

5'-most EST LIB3050-005-Q1-K1-E2

Method BLASTX
NCBI GI g2981475
BLAST score 1301
E value 1.0e-144
Match length 325
% identity 74

% identity 74
NCBI Description (AF053084) putative cinnamyl alcohol dehydrogenase [Malus

domestica]

Seq. No. 11067

Contig ID 20447 1.R1040

5'-most EST LIB3050-012-Q1-E1-D6

Seq. No. 11068

Contig ID 20450 1.R1040

5'-most EST uC-gmrominsoy042b07b1

Seq. No. 11069

Contig ID 20452_1.R1040

5'-most EST LIB3139-036-P1-N1-G10

Seq. No. 11070

Contig ID 20453_1.R1040

5'-most EST LIB3139-096-P1-N1-H4

Seq. No. 11071

Contig ID 20453_2.R1040

5'-most EST LIB3072-024-Q1-E1-F2

Seq. No. 11072

Contig ID 20453_5.R1040

5'-most EST LIB3073-001-Q1-K1-H3

Seq. No. 11073

Contig ID 20453_6.R1040 5'-most EST uC-gmropic004e03b1

Seq. No. 11074

Contig ID 20458 1.R1040

5'-most EST LIB3051-083-Q1-K1-A4

Seq. No. 11075
Contig ID 20458_2.R1040
5'-most EST LIB3051-031-Q1-K1-C2
Method BLASTX
NCBI GI g285741

BLAST score 720 E value 9.0e-76 Match length 405 % identity 42

NCBI Description (D14550) EDGP precursor [Daucus carota]

Seq. No. 11076

Contig ID 20458 3.R1040

5'-most EST LIB3051-086-Q1-K1-B6

Seq. No. 11077

Contig ID 20461 1.R1040

5'-most EST LIB3050-012-Q1-E1-B6

Seq. No. 11078

Contig ID 20464 1.R1040

5'-most EST LIB3092-033-Q1-K1-A12

Method BLASTX
NCBI GI g4490736
BLAST score 257
E value 3.0e-22
Match length 80
% identity 34

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 11079

Contig ID 20464_2.R1040 5'-most EST gsv701048836.h1

Method BLASTX
NCBI GI 94490736
BLAST score 416
E value 9.0e-41
Match length 166
% identity 34

NCBI Description (AL035708) putative protein [Arabidopsis thaliana]

Seq. No. 11080

Contig ID 20466 1.R1040

5'-most EST LIB3050-012-Q1-E1-C10

Seq. No. 11081

Contig ID 20470_1.R1040 5'-most EST fua701039064.h1

Method BLASTX
NCBI GI g4567228
BLAST score 480
E value 5.0e-48
Match length 143
% identity 65

NCBI Description (AC007119) unknown protein [Arabidopsis thaliana]

```
11082
Seq. No.
                  20479 1.R1040
Contig ID
                  LIB3050-011-Q1-E1-H9
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2832623
BLAST score
                  703
                   5.0e-74
E value
                   207
Match length
% identity
                   (AL021711) protein kinase - like protein [Arabidopsis
NCBI Description
                   thaliana]
                   11083
Seq. No.
                   20481 1.R1040
Contig ID
                   gsv701056625.h1
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3953458
BLAST score
                   948
                   1.0e-103
E value
                   215
Match length
% identity
                   88
                   (AC002328) F20N2.3 [Arabidopsis thaliana]
NCBI Description
                   11084
Seq. No.
                   20482 1.R1040
Contig ID
                   k11701213506.hl
5'-most EST
Method
                   BLASTX
                   g4539301
NCBI GI
                   822
BLAST score
E value
                   8.0e-88
Match length
                   355
                   47
% identity
```

(AL049480) putative mitochondrial protein [Arabidopsis NCBI Description

thaliana]

11085 Seq. No.

20487 1.R1040 Contig ID $jsh70\overline{1}063973.h1$ 5'-most EST

11086 Seq. No.

20488_1.R1040 Contig ID

LIB3107-014-Q1-K1-F2 5'-most EST

BLASTX Method q4049350 NCBI GI 538 BLAST score 1.0e-54 E value 220 Match length 50 % identity

(AL034567) putative protein [Arabidopsis thaliana] NCBI Description

11087 Seq. No.

20497 1.R1040 Contig ID

LIB3109-038-Q1-K1-G8 5'-most EST

11088 Seq. No.

20500 1.R1040 Contig ID

LIB3139-047-P1-N1-B4 5'-most EST

NCBI Description



```
BLASTX
Method
                   g3309082
NCBI GI
                   656
BLAST score
                   1.0e-68
E value
                   164
Match length
% identity
                   77
                   (AF076251) calcineurin B-like protein 1 [Arabidopsis
NCBI Description
                   thaliana]
                   11089
Seq. No.
                   20501 1.R1040
Contig ID
                   dpv701102737.h1
5'-most EST
                   BLASTX
Method
                   q2598589
NCBI GI
                   853
BLAST score
                   1.0e-109
E value
Match length
                   275
                   69
% identity
                   (Y15367) MtN19 [Medicago truncatula]
NCBI Description
                   11090
Seq. No.
Contig ID
                   20501 2.R1040
                   LIB3050-011-Q1-E1-D7
5'-most EST
Method
                   BLASTX
                   g2598589
NCBI GI
                   590
BLAST score
                   5.0e-61
E value
                   177
Match length
% identity
                   (Y15367) MtN19 [Medicago truncatula]
NCBI Description
                   11091
Seq. No.
                   20511 1.R1040
Contig ID
                   LIB30\overline{5}0-011-Q1-E1-H4
5'-most EST
Method
                   BLASTX
                   q4406789
NCBI GI
BLAST score
                   433
                   9.0e-43
E value
                   114
Match length
% identity
                   73
                    (AC006532) putative glutamate decarboxylase [Arabidopsis
NCBI Description
                   thaliana]
                   11092
Seq. No.
                   20513 1.R1040
Contig ID
                   LIB3050-011-Q1-E1-G10
 5'-most EST
                    11093
 Seq. No.
                    20522 1.R1040
 Contig ID
                    LIB3107-049-Q1-K1-H5
 5'-most EST
                    BLASTX
 Method
                    q2244924
 NCBI GI
                    335
 BLAST score
                    2.0e-31
 E value
                    87
 Match length
                    72
 % identity
```

(Z97339) glutaredoxin [Arabidopsis thaliana]

```
11094
Seq. No.
                   20522 2.R1040
Contig ID
                  LIB3050-011-Q1-E1-F1
5'-most EST
                   BLASTX
Method
                   g2244924
NCBI GI
BLAST score
                   291
                   3.0e-26
E value
                   75
Match length
                   71
% identity
NCBI Description (Z97339) glutaredoxin [Arabidopsis thaliana]
                   11095
Seq. No.
                   20529 1.R1040
Contig ID
                   LIB3051-017-Q1-E1-C8
5'-most EST
                   BLASTX
Method
                   g4538913
NCBI GI
                   490
BLAST score
                   4.0e-49
E value
                   207
Match length
                   56
% identity
NCBI Description (AL049482) putative protein [Arabidopsis thaliana]
                   11096
Seq. No.
                   20532 1.R1040
Contig ID
                   g4293043
5'-most EST
                   BLASTX
Method
                   g417148
NCBI GI
                   695
BLAST score
                   5.0e-73
E value
                   215
Match length
                   63
% identity
                   PROBABLE GLUTATHIONE S-TRANSFERASE (HEAT SHOCK PROTEIN 26A)
NCBI Description
                    (G2-4) >gi 99912_pir A33654 heat shock protein 26A -
                   soybean > gi_169981 (\overline{M2}0363) Gmhsp26-A [Glycine max]
                   11097
Seq. No.
                   20538 1.R1040
Contig ID
                   LIB3050-011-Q1-E1-D12
 5'-most EST
                   BLASTX
Method
                    q2275204
NCBI GI
BLAST score
                    371
                    4.0e-35
 E value
```

Match length 104 % identity

(ACO02337) DNA binding protein isolog [Arabidopsis NCBI Description

thaliana]

Seq. No.

11098

Contig ID

20539 1.R1040

LIB3050-011-Q1-E1-D4 5'-most EST

Seq. No.

11099

20539 2.R1040 Contig ID

5'-most EST

jC-gmro02800032c01a1

Seq. No.

11100



20540 1.R1040 Contig ID 5'-most EST LIB3050-011-Q1-E1-D5

Seq. No. 11101

Contig ID 20543 1.R1040

5'-most EST jC-gmle01810037b02a2

BLASTX Method NCBI GI q112925 BLAST score 335 E value 6.0e-31 Match length 177 % identity 45

41-2 PROTEIN ANTIGEN PRECURSOR >gi 320938 pir A45503 41-2 NCBI Description

> protein antigen precursor - Plasmodium falciparum >qi 160039 (J04656) 41-2 protein antigen [Plasmodium

falciparum]

11102 Seq. No.

20544 1.R1040 Contig ID

5'-most EST LIB3050-011-Q1-E1-D9

Method BLASTX g2580584 NCBI GI BLAST score 625 E value 4.0e-65 Match length 208 % identity 59

NCBI Description (AF000976) 7-O-methyltransferase [Medicago sativa]

Seq. No. 11103

Contig ID 20545 1.R1040

5'-most EST LIB3092-047-Q1-K1-C3

Method BLASTX NCBI GI q1170620 BLAST score 701 6.0e-74E value 186 Match length 73 % identity

KINESIN-LIKE PROTEIN B >gi 1084341 pir S48019 NCBI Description

kinesin-related protein katB - Arabidopsis thaliana

>gi_1438842_dbj_BAA04673_ (D21137) heavy chain polypeptide

of kinesin-like protein [Arabidopsis thaliana]

Seq. No. 11104

Contig ID 20545 2.R1040

5'-most EST LIB3050-011-Q1-E1-E11

Method BLASTX NCBI GI g1170619 BLAST score 557 E value 2.0e-57 Match length 149 76 % identity

NCBI Description KINESIN-LIKE PROTEIN A >gi 479594 pir S34830

kinesin-related protein katA - Arabidopsis thaliana >gi_303502_dbj_BAA01972_ (D11371) kinesin-like motor

protein heavy chain [Arabidopsis thaliana]

>gi 2911084 emb CAA17546 (AL021960) kinesin-related protein katA [Arabidopsis thaliana]

% identity

Seq. No.

Contig ID

NCBI Description

48

11111

20572 2.R1040



```
11105
Seq. No.
                   20562 1-R1040
Contig ID
                   jC-gmst02400016e09a1
5'-most EST
                   BLASTX
Method
                   g4309698
NCBI GI
                   227
BLAST score
                   3.0e-18
E value
                   63
Match length
                   78
% identity
                   (AC006266) putative glucosyltransferase [Arabidopsis
NCBI Description
                   thaliana]
                   11106
Seq. No.
                   20563 1.R1040
Contig ID
                   LIB3050-011-Q1-E1-A8
5'-most EST
                   BLASTX
Method
                   g1362018
NCBI GI
                   189
BLAST score
                   4.0e-14
E value
                   88
Match length
                   53
% identity
                   zinc finger protein 5 - Arabidopsis thaliana >gi_790681
NCBI Description
                   (L39648) zinc finger protein [Arabidopsis thaliana]
                   11107
Seq. No.
                   20564 2.R1040
Contig ID
                   jC-gmle01810031f07d1
5'-most EST
                   11108
Seq. No.
                   20565 1.R1040
Contig ID
                   LIB3051-015-Q1-E1-G1
5'-most EST
Method
                   BLASTX
                   q1762933
NCBI GI
BLAST score
                   267
E value
                   3.0e-23
                   132
Match length
% identity
                   (U66263) tumor-related protein [Nicotiana tabacum]
NCBI Description
                   11109
Seq. No.
                   20570 1.R1040
Contig ID
5'-most EST
                   LIB3109-055-Q1-K1-E11
Seq. No.
                   11110
                   20572 1.R1040
Contig ID
                   uC-gmropic073f05b1
5'-most EST
Method
                   BLASTX
                   g4432841
NCBI GI
                   555
BLAST score
                   1.0e-68
E value
                   384
Match length
```

(AC006283) hypothetical protein [Arabidopsis thaliana]



```
zsg701117769.h2
5'-most EST
Method
                  BLASTX
NCBI GI
                  g4432841
BLAST score
                  296
E value
                  2.0e-26
Match length
                  97
% identity
                  (AC006283) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  11112
Seq. No.
                  20580 1.R1040
Contig ID
                  uC-gmropic009b03b1
5'-most EST
                  BLASTX
Method
                  q3695389
NCBI GI
                  351
BLAST score
E value
                   4.0e-33
Match length
                  127
                   27
% identity
                   (AF096371) contains similarity to D-isomer specific
NCBI Description
                   2-hydroxyacid dehydrogenases (Pfam: 2-Hacid_DH.hmm, score:
                  19.11) [Arabidopsis thaliana]
                   11113
Seq. No.
                   20580 2.R1040
Contig ID
                  uC-gmronoir015e07b1
5'-most EST
Seq. No.
                   11114
                   20581 1.R1040
Contig ID
                   q5753648
5'-most EST
Method
                   BLASTX
                   g1732556
NCBI GI
BLAST score
                   198
E value
                   3.0e-15
Match length
                   91
% identity
                   59
                  (U66316) desiccation protective protein LEA5 [Glycine max]
NCBI Description
Seq. No.
                   11115
                   20581 2.R1040
Contig ID
                   LIB3050-010-Q1-E1-H12
5'-most EST
Method
                   BLASTX
                   q1084390
NCBI GI
                   202
BLAST score
                   1.0e-15
E value
                   88
Match length
                   48
% identity
                   late-embryogenesis protein lea5 - sweet orange
NCBI Description
                   >gi 599726 emb CAA86851_ (Z46824) Lea5 protein [Citrus
                   sinensis]
                   11116
Seq. No.
                   20588 1.R1040
Contig ID
                   LIB3050-010-Q1-E1-G12
5'-most EST
```

jC-gmle01810068b12d1

1926

11117

20591 1.R1040

Seq. No.

Contig ID 5'-most EST



11118 Seq. No. 20591 2.R1040 Contig ID jC-gmle01810071b09a1 5'-most EST 11119 Seq. No. 20595 1.R1040 Contig ID pmv700889116.h1 5'-most EST BLASTX Method q4490342 NCBI GI BLAST score 617 4.0e-64 E value 216 Match length 64 % identity (AL035656) putative protein [Arabidopsis thaliana] NCBI Description 11120 Seq. No. 20597 1.R1040 Contig ID trc700565580.h1 5'-most EST 11121 Seq. No. 20599 1.R1040 Contig ID hyd700724906.h1 5'-most EST BLASTX Method q4056490 NCBI GI 512 BLAST score 1.0e-51 E value 276 Match length 42 % identity (AC005896) hypothetical protein [Arabidopsis thaliana] NCBI Description 11122 Seq. No. 20599 2.R1040 Contig ID uC-qmflminsoy075c06b1 5'-most EST 11123 Seq. No. 20599 4.R1040 Contig ID kmv700741926.h1 5'-most EST BLASTX Method g4056490 NCBI GI BLAST score 712 3.0e-75 E value Match length 264 56 % identity (AC005896) hypothetical protein [Arabidopsis thaliana] NCBI Description Seq. No. 11124 20599 6.R1040 Contig ID vwf700675878.h1 5'-most EST

11125 Seq. No.

20602_1.R1040 Contig ID

LIB3050-010-Q1-E1-D9 5'-most EST

11126 Seq. No.

20605 1.R1040 Contig ID

LIB3107-069-Q1-K1-A8 5'-most EST



```
11127
Seq. No.
Contig ID
                   20609 1.R1040
5'-most EST
                  g5677800
Method
                  BLASTX
NCBI GI
                   g1146165
BLAST score
                   626
                   6.0e-65
E value
                   154
Match length
                   81
% identity
                  (L47479) urophorphyrin III methylase [Arabidopsis thaliana]
NCBI Description
                   11128
Seq. No.
                   20610 1.R1040
Contig ID
5'-most EST
                   ncj70\overline{0}985028.h1
Method
                   BLASTX
NCBI GI
                   q3122858
BLAST score
                   1283
                   1.0e-142
E value
                   307
Match length
                   83
% identity
                   D-3-PHOSPHOGLYCERATE DEHYDROGENASE PRECURSOR
NCBI Description
                   >gi_2189964_dbj_BAA20405 (AB003280) Phosphoglycerate
                   dehydrogenase [Arabidopsis thaliana]
                   >gi_2804258_dbj_BAA24440_ (AB010407) phosphoglycerate
                   dehydrogenase [Arabidopsis thaliana]
                   11129
Seq. No.
Contig ID
                   20622 1.R1040
                   zsq701117761.h2
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3132696
                   2115
BLAST score
                   0.0e+00
E value
                   462
Match length
                   91
% identity
                   (AF061962) SAR DNA-binding protein-1 [Pisum sativum]
NCBI Description
                   11130
Seq. No.
                   20628 1.R1040
Contig ID
                   LIB3093-046-Q1-K1-B5
5'-most EST
Seq. No.
                   11131
                   20628 2.R1040
Contig ID
                   LIB3087-002-Q1-K1-F10
5'-most EST
Seq. No.
                   11132
                   20628 3.R1040
Contig ID
                   LIB3139-113-P1-N1-D4
5'-most EST
                   11133
 Seq. No.
Contig ID
                   20633 1.R1040
```

Seq. No. 11134
Contig ID 20636_1.R1040
5'-most EST pmv700891688.h1



```
BLASTX
Method
                  q1707021
NCBI GI
BLAST score
                  450
                  5.0e-45
E value
Match length
                  109
% identity
                  72
                   (U78721) Ubiquitin-conjugating enzyme, E2-16kD isolog
NCBI Description
                   [Arabidopsis thaliana]
                  11135
Seq. No.
                  20642 1.R1040
Contig ID
                  LIB3050-010-Q1-E1-A2
5'-most EST
                  BLASTX
Method
                  g4510398
NCBI GI
BLAST score
                   231
E value
                   5.0e-19
                   111
Match length
% identity
                   52
                  (AC006587) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   11136
Seq. No.
                   20643 1.R1040
Contig ID
                   LIB3050-010-Q1-E1-A3
5'-most EST
                   BLASTX
Method
                   g2565010
NCBI GI
BLAST score
                   472
E value
                   3.0e-47
                   108
Match length
% identity
                   (AC002983) putative microfibril-associated protein
NCBI Description
                   [Arabidopsis thaliana] >gi_3377811 (AF076275) contains
                   similarity to ATP synthase B/B' (Pfam: ATP-synt_B.hmm,
                   score: 11.71) [Arabidopsis thaliana]
                   11137
Seq. No.
                   20645 1.R1040
Contig ID
                   hrw701062847.h1
5'-most EST
                   BLASTX
Method
                   g2129597
NCBI GI
                   1954
BLAST score
                   0.0e + 00
E value
Match length
                   411
 % identity
                   89
                   glutamate dehydrogenase 1 - Arabidopsis thaliana
NCBI Description
                   >gi 1098960 (U37771) glutamate dehydrogenase 1 [Arabidopsis
                   thaliana] >gi_1293095 (U53527) glutamate dehydrogenase 1
                   [Arabidopsis Thaliana]
                   11138
 Seq. No.
                   20645 2.R1040
 Contig ID
                   fC-qmf1700906513f4
 5'-most EST
                   BLASTX
 Method
```

5'-most EST fC-gmf1700
Method BLASTX
NCBI GI g2129597
BLAST score 678
E value 2.0e-71
Match length 152
% identity 88



glutamate dehydrogenase 1 - Arabidopsis thaliana NCBI Description >gi 1098960 (U37771) glutamate dehydrogenase 1 [Arabidopsis thaliana] >gi 1293095 (U53527) glutamate dehydrogenase 1 [Arabidopsis thaliana] 11139 Seq. No. 20645 3.R1040 Contig ID LIB3109-035-Q1-K6-D10 5'-most EST Method BLASTX g2129597 NCBI GI BLAST score 332 1.0e-30 E value Match length 72 90 % identity glutamate dehydrogenase 1 - Arabidopsis thaliana NCBI Description >gi 1098960 (U37771) glutamate dehydrogenase 1 [Arabidopsis thaliana] >gi_1293095 (U53527) glutamate dehydrogenase 1 [Arabidopsis thaliana] 11140 Seq. No. Contig ID 20655 1.R1040 5'-most EST leu701149415.hl Method BLASTX q3157934 NCBI GI BLAST score 611 E value 1.0e-63164 Match length 72 % identity (AC002131) Similar to hypothetical protein F09E5.8 NCBI Description gb U37429 from C. elegans. ESTs gb_T42019 and gb_N97000 come from this gene. [Arabidopsis thaliana] 11141 Seq. No. 20670 1.R1040 Contig ID LIB3094-073-Q1-K1-E8 5'-most EST BLASTX Method g4454457 NCBI GI 466 BLAST score 2.0e-46 E value 116 Match length % identity (AC006234) unknown protein [Arabidopsis thaliana] NCBI Description 11142 Seq. No. 20670 2.R1040 Contig ID 5'-most EST pmv700892880.h1 Method BLASTX q4454457 NCBI GI 370

Method BLASTX
NCBI GI 94454457
BLAST score 370
E value 2.0e-35
Match length 84
% identity 85

NCBI Description (AC006234) unknown protein [Arabidopsis thaliana]

Seq. No.

11143

Contig ID 20670_5.R1040

5'-most EST LIB3094-031-Q1-K1-C5



```
BLASTN
Method
                   q4454447
NCBI GI
BLAST score
                   52
                   3.0e-20
E value
Match length
                   100
% identity
                   88
                   Arabidopsis thaliana chromosome II BAC F5H14 genomic
NCBI Description
                   sequence, complete sequence [Arabidopsis thaliana]
                   11144
Seq. No.
Contig ID
                   20672 1.R1040
                   LIB3050-009-Q1-E1-H5
5'-most EST
                   BLASTX
Method
                   g1707016
NCBI GI
                   144
BLAST score
E value
                   5.0e-09
Match length
                   111
                   33
% identity
                   (U78721) cadmium-induced protein isolog [Arabidopsis
NCBI Description
                   thaliana]
                   11145
Seq. No.
                   20673 1.R1040
Contig ID
                   zzp700832934.h1
5'-most EST
                   11146
Seq. No.
Contig ID
                   20679 1.R1040
                   LIB3050-009-Q1-E1-E2
5'-most EST
                   11147
Seq. No.
                   20679 2.R1040
Contig ID
5'-most EST
                   pmv700890721.h1
                   11148
Seq. No.
                   20684 1.R1040
Contig ID
                   epx70\overline{1}103717.h1
5'-most EST
Method
                   BLASTX
                   g3004565
NCBI GI
                   207
BLAST score
                   7.0e-16
E value
Match length
                   193
% identity
                   15
                   (AC003673) putative protein kinase [Arabidopsis thaliana]
NCBI Description
                   11149
Seq. No.
                   20685 1.R1040
Contig ID
                   ssr700556138.hl
5'-most EST
                   11150
Seq. No.
                    20694 1.R1040
Contig ID
                    jC-gmle01810055b11a1
 5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g3068717
BLAST score 698
E value 1.0e-81

Match length 220 % identity 68



```
(AF049236) unknown [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  11151
Contig ID
                  20695 1.R1040
                  LIB3050-009-Q1-E1-C4
5'-most EST
                  BLASTX
Method
                  g4580393
NCBI GI
BLAST score
                   412
                   3.0e-40
E value
Match length
                  155
% identity
                   51
                  (AC007171) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   11152
Seq. No.
Contig ID
                   20712 1.R1040
                   jC-gmst02400050g04a1
5'-most EST
Method
                   BLASTX
                   g3786011
NCBI GI
BLAST score
                   1069
E value
                   1.0e-117
                   239
Match length
% identity
                   88
                   (AC005499) putative elongation factor [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   11153
                   20712 2.R1040
Contig ID
                   LIB3050-009-Q1-E1-A3
5'-most EST
Method
                   BLASTX
                   q3786011
NCBI GI
                   423
BLAST score
E value
                   1.0e-41
                   98
Match length
                   89
% identity
                   (AC005499) putative elongation factor [Arabidopsis
NCBI Description
                   thaliana]
                   11154
Seq. No.
                   20712 3.R1040
Contig ID
                   jC-gmro02800030d09a1
5'-most EST
                   BLASTX
Method
                   g3786011
NCBI GI
                   264
BLAST score
                   5.0e-33
E value
                   94
Match length
                   84
% identity
                   (AC005499) putative elongation factor [Arabidopsis
NCBI Description
                   thaliana]
                   11155
Seq. No.
Contig ID
                   20712 4.R1040
                   ncj700978495.hl
5'-most EST
                   BLASTX
Method
                   q3786011
NCBI GI
```

1932

264

66

2.0e-23

BLAST score E value

Match length



Seq. No. 11156

Contig ID 20713_1.R1040

5'-most EST jC-gmle01810006g11a1

Method BLASTX
NCBI GI g3482979
BLAST score 149
E value 6.0e-09
Match length 206
% identity 28

NCBI Description (AL031369) putative protein [Arabidopsis thaliana]

>gi_4567258_gb_AAD23672.1_AC007070_21 (AC007070)

hypothetical protein [Arabidopsis thaliana]

Seq. No. 11157

Contig ID 20713_2.R1040

5'-most EST LIB3170-062-Q1-J1-E3

Seq. No. 11158

Contig ID 20713_3.R1040 5'-most EST epx701108675.h1

Seq. No. 11159

Contig ID 20715_1.R1040 5'-most EST uC-gmropic091h07b1

Seq. No. 11160

Contig ID 20738_1.R1040 5'-most EST pmv700894377.h1

Method BLASTX
NCBI GI g2289006
BLAST score 703
E value 6.0e-74
Match length 180
% identity 73

NCBI Description (AC002335) glutathione perosidase isolog [Arabidopsis

thaliana]

Seq. No. 11161

Contig ID 20738_2.R1040 5'-most EST pxt700946407.h1

Seq. No. 11162

Contig ID 20738 3.R1040

5'-most EST LIB3170-049-Q1-K2-E6

Seq. No. 11163

Contig ID 20738_4.R1040

5'-most EST jC-gmf102220108d04a1

Method BLASTX
NCBI GI g1946690
BLAST score 526
E value 1.0e-53
Match length 117

```
% identity
                    (U94495) glutathione peroxidase [Arabidopsis thaliana]
 NCBI Description
 Seq. No.
                    11164
                    20739 1.R1040
 Contig ID
                    LIB30\overline{5}0-008-Q1-E1-H4
 5'-most EST
                    BLASTX
 Method
                    q1699024
 NCBI GI
                    727
 BLAST score
                    9.0e-77
. E value
                    233
 Match length
                    58
 % identity
                    (U78866) gene1000 [Arabidopsis thaliana] >gi_1699057
 NCBI Description
                     (U78870) unknown [Arabidopsis thaliana]
 Seq. No.
                    11165
                    20739 2.R1040
 Contig ID
                    leu70\overline{1}146127.h1
 5'-most EST
                    BLASTX
 Method
                     q1699024
 NCBI GI
                     350
 BLAST score
                     4.0e-33
 E value
                     100
 Match length
 % identity
                     65
                     (U78866) gene1000 [Arabidopsis thaliana] >gi_1699057
 NCBI Description
                     (U78870) unknown [Arabidopsis thaliana]
 Seq. No.
                     11166
                     20739 3.R1040
 Contig ID
                     jC-gmst02400039a04a1
 5'-most EST
                     BLASTX
 Method
                     g1699024
 NCBI GI
                     269
 BLAST score
                     2.0e-23
 E value
                     150
 Match length
  % identity
                     43
                    (U78866) gene1000 [Arabidopsis thaliana] >gi_1699057
 NCBI Description
                     (U78870) unknown [Arabidopsis thaliana]
                     11167
  Seq. No.
                     20739 5.R1040
  Contig ID
                     eep70\overline{0}863855.h1
  5'-most EST
                     11168
  Seq. No.
                     20740 1.R1040
  Contig ID
```

trc700565610.hl 5'-most EST

BLASTX Method q1708424 NCBI GI BLAST score 1036 1.0e-113 E value Match length 262 76 % identity

NCBI Description ISOFLAVONE REDUCTASE HOMOLOG >gi_1230614 (U48590) isoflavone reductase-like protein [Lupinus albus]

11169 Seq. No.

20741 1.R1040 Contig ID



```
jC-gmf102220114a10d1
5'-most EST
                  11170
Seq. No.
                  20749 1.R1040
Contig ID
                  LIB3050-008-Q1-E1-F11
5'-most EST
                  BLASTN
Method
                  q4432811
NCBI GI
BLAST score
                  61
                  3.0e-25
E value
                  125
Match length
% identity
                  87
                  Arabidopsis thaliana chromosome II BAC F16D14 genomic
NCBI Description
                  sequence, complete sequence [Arabidopsis thaliana]
                  11171
Seq. No.
Contig ID
                   20754 1.R1040
                  LIB3050-008-Q1-E1-F6
5'-most EST
                   11172
Seq. No.
                   20758 1.R1040
Contig ID
                   jC-gmst02400049c04a1
5'-most EST
Seq. No.
                   11173
                   20758 3.R1040
Contig ID
                   LIB3050-008-Q1-E1-G1
5'-most EST
Seq. No.
                   11174
                   20760 1.R1040
Contig ID
                   LIB3107-041-Q1-K1-F4
5'-most EST
Method
                   BLASTN
                   g667048
NCBI GI
BLAST score
                   377
                   0.0e + 00
E value
                   536
Match length
                   93
 % identity
NCBI Description P.vulgaris mRNA for protein phosphatase 1 (PP1)
                   11175
 Seq. No.
                   20760 2.R1040
 Contig ID
                   LIB3167-013-P1-K1-D11
 5'-most EST
                   BLASTN
 Method
                   g667048
 NCBI GI
                   178
 BLAST score
                   2.0e-95
 E value
                   266
 Match length
                   92
 % identity
NCBI Description P.vulgaris mRNA for protein phosphatase 1 (PP1)
                   11176
 Seq. No.
                   20760 3.R1040
 Contig ID
                    fua701038292.hl
 5'-most EST
 Method
                    BLASTN
                    q667048
 NCBI GI
                    109
 BLAST score
                    3.0e-54
 E value
                    181
 Match length
```

1935

90

% identity



```
NCBI Description P.vulgaris mRNA for protein phosphatase 1 (PP1)
                  11177
Seq. No.
Contig ID
                  20763 1.R1040
5'-most EST
                  LIB3050-008-Q1-E1-E7
Method
                  BLASTX
                  g1729938
NCBI GI
BLAST score
                  321
                  3.0e-29
E value
Match length
                  101
% identity
                  56
                  THIOREDOXIN-LIKE PROTEIN SLR0233 >gi_1001610_dbj_BAA10238_
NCBI Description
                   (D64000) thioredoxin M [Synechocystis sp.]
                  11178
Seq. No.
Contig ID
                  20763 2.R1040
5'-most EST
                  LIB3170-002-Q1-K1-A2
                  11179
Seq. No.
Contig ID
                   20765 1.R1040
5'-most EST
                  jC-qmro02910034a10d1
                   11180
Seq. No.
Contig ID
                   20770 1.R1040
5'-most EST
                  pxt700945188.hl
Seq. No.
                   11181
Contig ID
                   20772 1.R1040
5'-most EST
                   LIB3094-006-Q1-K1-C1
Method
                   BLASTX
NCBI GI
                   g4039014
BLAST score
                   783
E value
                   1.0e-106
Match length
                   372
                   51
% identity
                  (AF037338) cleft lip and palate transmembrane protein 1
NCBI Description
                   [Homo sapiens] >gi_4063033 (AF037339) cleft lip and palate
                   transmembrane protein 1 [Homo sapiens]
                   >gi 4502897 ref NP 001285.1 pCLPTM1 cleft lip and palate
                   associated transmembrane protein
Seq. No.
                   11182
                   20787 1.R1040
Contig ID
                   LIB3050-008-Q1-E1-C5
5'-most EST
                   11183
Seq. No.
                   20790 1.R1040
Contig ID
                   ncj700980962.hl
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3123295
```

Method BLASTX
NCBI GI g3123295
BLAST score 248
E value 1.0e-20
Match length 136
% identity 17

NCBI Description CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi_2583169 (AF026473) calmodulin-related protein [Arabidopsis

thaliana]



Seq. No. 11184

Contig ID 20790_2.R1040 5'-most EST zzp700832902.h1

Seq. No. 11185

Contig ID 20794 1.R1040

5'-most EST jC-gmst02400069d05a1

Method BLASTX
NCBI GI g4206196
BLAST score 316
E value 1.0e-28
Match length 153
% identity 42

NCBI Description (AF071527) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11186

Contig ID 20795_1.R1040 5'-most EST uxk700667764.h1

Method BLASTX
NCBI GI g2827621
BLAST score 157
E value 3.0e-10
Match length 143
% identity 30

NCBI Description (AL021636) putative protein [Arabidopsis thaliana]

Seq. No. 11187

Contig ID 20804_1.R1040 5'-most EST uC-gmropic057b03b1

Method BLASTX
NCBI GI g3738320
BLAST score 672
E value 3.0e-70
Match length 282
% identity 47

NCBI Description (AC005170) putative cinnamoyl CoA reductase [Arabidopsis

thaliana]

Seq. No. 11188

Contig ID 20804_2.R1040 5'-most EST fua701041128.h1

Method BLASTX
NCBI GI g3738320
BLAST score 184
E value 2.0e-13
Match length 95
% identity 41

NCBI Description (AC005170) putative cinnamoyl CoA reductase [Arabidopsis

thaliana]

Seq. No. 11189

Contig ID 20814 1.R1040 5'-most EST sat701012928.h1

Method BLASTN
NCBI GI g4530125
BLAST score 200

```
1.0e-108
E value
                  384
Match length
% identity
                  88
                  Phaseolus vulgaris receptor-like protein kinase homólog
NCBI Description
                  RK20-1 mRNA, complete cds
                  11190
Seq. No.
                  20817 1.R1040
Contia ID
                  LIB3050-007-Q1-E1-G4
5'-most EST
                  BLASTX
Method
                  g3004556
NCBI GI
BLAST score
                  246
                  6.0e-21
E value
                  93
Match length
% identity
                  (AC003673) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  11191
                  20825 1.R1040
Contig ID
                  sat701011326.h1
5'-most EST
                  BLASTX
Method
NCBI GI
                   q3549679
BLAST score
                   752
                   1.0e-79
E value
                   259
Match length
                   62
% identity
                  (AL031394) putative protein [Arabidopsis thaliana]
NCBI Description
                   11192
Seq. No.
                   20825 2.R1040
Contig ID
                   LIB3092-030-Q1-K1-D4
5'-most EST
                   11193
Seq. No.
                   20825 3.R1040
Contig ID
                   epx701105311.hl
5'-most EST
                   BLASTX
Method
                   g3549679
NCBI GI
                   398
BLAST score
                   3.0e-38
E value
                   176
Match length
                   52
% identity
                  (AL031394) putative protein [Arabidopsis thaliana]
NCBI Description
                   11194
Seq. No.
                   20825 5.R1040
Contig ID
                   LIB3056-014-Q1-N1-F2
5'-most EST
                   11195
Seq. No.
                   20825 6.R1040
Contig ID
                   uC-gmrominsoy097g08b1
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3549679
                   360
BLAST score
```

3.0e-34 E value 132 Match length 59 % identity

NCBI Description (AL031394) putative protein [Arabidopsis thaliana]



```
11196
Seq. No.
                  20825 7.R1040
Contig ID
                  ssr70\overline{0}560602.h1
5'-most EST
Method
                  BLASTX
                  g3549679
NCBI GI
                  170
BLAST score
                  7.0e-12
E value
Match length
                  61
                  57
% identity
NCBI Description (AL031394) putative protein [Arabidopsis thaliana]
                   11197
Seq. No.
                   20825 8.R1040
Contig ID
                   LIB3051-003-Q1-E1-G7
5'-most EST
Method
                   BLASTX
                   g3549679
NCBI GI
BLAST score
                   167
                   1.0e-11
E value
                   86
Match length
                   60
% identity
                  (AL031394) putative protein [Arabidopsis thaliana]
NCBI Description
                   11198
Seq. No.
                   20828 1.R1040
Contig ID
                   LIB3050-007-Q1-E1-F7
5'-most EST
Method
                   BLASTX
NCBI GI
                   g2781355
BLAST score
                   284
                   2.0e-25
E value
                   106
Match length
% identity
                   56
                   (AC003113) F2401.11 [Arabidopsis thaliana]
NCBI Description
                   11199
Seq. No.
                   20832 1.R1040
Contig ID
                   LIB3050-005-Q1-K1-E4
5'-most EST
                   11200
Seq. No.
                   20833 1.R1040
Contig ID
5'-most EST
                   zzp700831448.h1
                   BLASTX
Method
                   q2244806
NCBI GI
                   568
BLAST score
                   6.0e-58
E value
                   296
Match length
                   46
% identity
                   (Z97336) hypothetical protein [Arabidopsis thaliana]
NCBI Description
                   11201
Seq. No.
                   20833_2.R1040
Contig ID
                   LIB3139-107-P1-N1-C1
5'-most EST
                   BLASTX
Method
                   g2244806
NCBI GI
                   355
BLAST score
```

3.0e - 33

113

E value Match length



% identity 58

NCBI Description (Z97336) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11202

Contig ID 20833_3.R1040 5'-most EST g5342537

Seq. No. 11203

Contig ID 20857_1.R1040 5'-most EST fC-gmle700685573f1

Method BLASTX
NCBI GI g3859603
BLAST score 311
E value 4.0e-32
Match length 188
% identity 47

NCBI Description (AF104919) contains similarity to wild oat DNA-binding

protein ABF2 (GB:Z48431) [Arabidopsis thaliana]

Seq. No. 11204

Contig ID 20860_1.R1040 5'-most EST dpv701097008.h1

Method BLASTX
NCBI GI g3123295
BLAST score 203
E value 1.0e-15
Match length 145
% identity 29

NCBI Description CALMODULIN-RELATED PROTEIN 2, TOUCH-INDUCED >gi_2583169

(AF026473) calmodulin-related protein [Arabidopsis

thaliana]

Seq. No. 11205

Contig ID 20862_1.R1040

5'-most EST LIB3170-037-Q1-K1-E12

Method BLASTX
NCBI GI g1707017
BLAST score 1332
E value 1.0e-148
Match length 284
% identity 89

NCBI Description (U78721) RNA helicase isolog [Arabidopsis thaliana]

Seq. No. 11206

Contig ID 20866 1.R1040

5'-most EST jC-gmro02910024g10a1

Method BLASTX
NCBI GI 94309759
BLAST score 1578
E value 1.0e-176
Match length 431
% identity 73

NCBI Description (AC006217) unknown protein with Src homology 3 (SH3) domain

profile (PDOC50002) [Arabidopsis thaliana]

Seq. No. 11207

Contig ID 20883_1.R1040

1940



```
g5057935
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3193296
BLAST score
                  959
E value
                  1.0e-104
Match length
                  230
% identity
                  77
                  (AF069298) similar to pectinesterase [Arabidopsis thaliana]
NCBI Description
                  11208
Seq. No.
                  20885 1.R1040
Contig ID
                  uC-gmronoir044d01b1
5'-most EST
                   11209
Seq. No.
                   20886 1.R1040
Contig ID
5'-most EST
                   zhf700952345.h1
Method
                   BLASTX
NCBI GI
                   g3335341
BLAST score
                   287
E value
                   2.0e-25
Match length
                   77
% identity
                   74
                  (AC004512) T8F5.10 [Arabidopsis thaliana]
NCBI Description
                   11210
Seq. No.
Contig ID
                   20887 1.R1040
                   LIB3050-007-Q1-E1-A10
5'-most EST
                   11211
Seq. No.
                   20894 1.R1040
Contig ID
                   zhf700963376.hl
5'-most EST
Method
                   BLASTX
                   q2647949
NCBI GI
BLAST score
                   516
                   6.0e-57
E value
                   130
Match length
% identity
                   73
                   (AJ001369) cytochrome b5 [Olea europaea]
NCBI Description
                   11212
Seq. No.
                   20898 1.R1040
Contig ID
                   LIB3050-006-Q1-E1-H4
5'-most EST
Method
                   BLASTX
                   q3142290
NCBI GI
                   633
BLAST score
                   6.0e-66
E value
                   196
Match length
                   68
% identity
                   (AC002411) Contains similarity to gb_Z69902 from C.
NCBI Description
                   elegans. [Arabidopsis thaliana]
                   11213
Seq. No.
                   20908 1.R1040
Contig ID
                   LIB3139-106-P1-N1-F8
 5'-most EST
```

BLASTX

1551

g2130073

Method

NCBI GI BLAST score



1.0e-173 E value 358 Match length % identity

fructose-bisphosphate aldolase (EC 4.1.2.13) isoenzyme C-1, cytosolic - rice >gi_786178_dbj_BAA08845_ (D50307) aldolase NCBI Description

C-1 [Oryza sativa] >gi_790970_dbj_BAA08830_ (D50301)

aldolase C-1 [Oryza sativa]

11214 Seq. No.

20911 1.R1040 Contig ID

LIB3050-006-Q1-E1-H10 5'-most EST

11215 Seq. No.

20912 1.R1040 Contig ID leu701146187.h1 5'-most EST

Method BLASTX NCBI GI q2245136 BLAST score 1349 1.0e-150 E value 333 Match length 73 % identity

(Z97344) trehalose-6-phosphate synthase homolog NCBI Description

[Arabidopsis thaliana]

11216 Seq. No.

20917 1.R1040 Contig ID

jC-qmro02910067f09d1 5'-most EST

11217 Seq. No.

Contig ID 20919 1.R1040 fC-gmse700851501a3 5'-most EST

Seq. No. 11218

20921 1.R1040 Contig ID

LIB3109-012-Q1-K1-H4 5'-most EST

Method BLASTN NCBI GI g3319883 BLAST score 35 6.0e-10 E value 77 Match length

% identity 95

NCBI Description Arabidopsis thaliana PRT1 gene

11219 Seq. No.

20922 1.R1040 Contig ID jex700908390.hl 5'-most EST

Seq. No. 11220

20931 1.R1040 Contig ID

LIB3050-006-Q1-E1-F8 5'-most EST

11221 Seq. No.

20941 1.R1040 Contig ID

LIB3050-006-Q1-E1-D7 5'-most EST

11222 Seq. No.

20943 1.R1040 Contig ID

E value

2.0e-56



```
wvk700684733.h1
5'-most EST
                  11223
Seq. No.
Contig ID
                  20947 1.R1040
5'-most EST
                  LIB3050-006-Q1-E1-E12
Method
                  BLASTX
                  q4006881
NCBI GI
BLAST score
                  654
                  4.0e-68
E value
                  194
Match length
% identity
                  66
NCBI Description (Z99707) putative protein [Arabidopsis thaliana]
Seq. No.
                  11224
                  20947 2.R1040
Contig ID
5'-most EST
                  LIB3106-040-Q1-K1-D6
Seq. No.
                  11225
                  20950 1.R1040
Contig ID
                  seb700653241.hl
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2104681
BLAST score
                  495
                  1.0e-49
E value
                  153
Match length
% identity
                   64
NCBI Description (X97907) transcription factor [Vicia faba]
                  11226
Seq. No.
Contig ID
                  20950 2.R1040
5'-most EST
                  LIB3050-006-Q1-E1-B4
                  BLASTN
Method
NCBI GI
                   g2104680
BLAST score
                   65
                   9.0e-28
E value
Match length
                   214
% identity
                   87
NCBI Description V.faba mRNA for putative transciption factor (1556bp)
Seq. No.
                   11227
Contig ID
                   20950 3.R1040
5'-most EST
                   LIB3051-045-Q1-K1-E8
Method
                   BLASTX
                   g2104681
NCBI GI
BLAST score
                   287
                   2.0e-25
E value
                   180
Match length
% identity
                   41
NCBI Description (X97907) transcription factor [Vicia faba]
                   11228
Seq. No.
Contig ID
                   20952 1.R1040
                   jC-gmst02400047e11a1
5'-most EST
                   BLASTX
Method
                   g627468
NCBI GI
BLAST score
                   382
```

1943



Match length 348 % identity 40

NCBI Description hypothetical protein 1 - human >gi_285983_dbj_BAA02799_

(D13635) KIAA0010 [Homo sapiens]

Seq. No. 11229

Contig ID 20952 2.R1040

5'-most EST LIB3050-006-Q1-E1-B6

Method BLASTX
NCBI GI g627468
BLAST score 324
E value 2.0e-35
Match length 159
% identity 49

NCBI Description hypothetical protein 1 - human >gi_285983_dbj_BAA02799_

(D13635) KIAA0010 [Homo sapiens]

Seq. No. 11230

Contig ID 20953_1.R1040 5'-most EST awf700836982.h1

Method BLASTX
NCBI GI g462567
BLAST score 638
E value 3.0e-75
Match length 388
% identity 45

NCBI Description MANNOSE-6-PHOSPHATE ISOMERASE (PHOSPHOMANNOSE ISOMERASE)

(PMI) (PHOSPHOHEXOMUTASE) >gi_631328_pir_S41122 mannose-6-phosphate isomerase (EC 5.3.1.8) - human

>gi_416017_emb_CAA53657_ (X76057) phosphomannose isomerase
[Homo sapiens] >gi_740967_prf__2006245A phosphomannose
isomerase [Homo sapiens] >gi_4505235_ref_NP_002426.1_pMPI_

mannose phosphate isomerase

Seq. No. 11231

Contig ID 20955_1.R1040 5'-most EST bth700848304.h1

Seq. No.

11232

Contig ID 20955 2.R1040

5'-most EST LIB3107-054-Q1-K1-H10

Seq. No.

11233

Contig ID 20956 1.R1040

5'-most EST LIB3050-006-Q1-E1-C1

Method BLASTX
NCBI GI g2245060
BLAST score 273
E value 4.0e-24
Match length 93
% identity 59

NCBI Description (Z97342) allergen homolog [Arabidopsis thaliana]

Seq. No. 11234

Contig ID 20960_1.R1040 5'-most EST g5676902 Method BLASTN



NCBI GI g4530125 BLAST score 120 E value 1.0e-60 Match length 466 % identity 83

NCBI Description Phaseolus vulgaris receptor-like protein kinase homolog

RK20-1 mRNA, complete cds

Seq. No. 11235

Contig ID 20960 2.R1040 g5605743 5'-most EST Method BLASTN g4530125 NCBI GI BLAST score 111 E value 2.0e-55 Match length 367 % identity 86

NCBI Description Phaseolus vulgaris receptor-like protein kinase homolog

RK20-1 mRNA, complete cds

Seq. No. 11236

Contig ID 20966_1.R1040 5'-most EST asj700967487.h1

Method BLASTX
NCBI GI g4185143
BLAST score 856
E value 6.0e-92
Match length 258
% identity 64

NCBI Description (AC005724) putative signal recognition particle receptor

beta subunit [Arabidopsis thaliana]

Seq. No. 11237

Contig ID 20966_2.R1040 5'-most EST zhf700962435.h1

Method BLASTX
NCBI GI g4185143
BLAST score 322
E value 1.0e-29
Match length 100
% identity 61

NCBI Description (AC005724) putative signal recognition particle receptor

beta subunit [Arabidopsis thaliana]

Seq. No. 11238

Contig ID 20968 1.R1040 5'-most EST g4301715

Seq. No. 11239

Contig ID 20968_2.R1040

5'-most EST LIB3050-006-Q1-E1-A12

Seq. No. 11240

Contig ID 20970 1.R1040

5'-most EST LIB3051-015-Q1-E1-B6

Method BLASTX NCBI GI g2829870



```
336
BLAST score
                  3.0e-31
E value
Match length
                  151
                  49
% identity
                  (AC002396) Hypothetical protein [Arabidopsis thaliana]
NCBI Description
                  11241
Seq. No.
Contig ID
                  20975 1.R1040
                  LIB3087-012-Q1-K1-B10
5'-most EST
                  BLASTX
Method
NCBI GI
                   g480618
BLAST score
                   650
                   1.0e-67
E value
                   234
Match length
% identity
                   ATAF1 protein - Arabidopsis thaliana (fragment)
NCBI Description
                   >gi 1345506_emb_CAA52771_ (X74755) ATAF1 [Arabidopsis
                   thaliana]
                   11242
Seq. No.
                   20975 2.R1040
Contig ID
                   pcp700989592.hl
5'-most EST
                   11243
Seq. No.
                   20996 1.R1040
Contig ID
                   LIB3050-005-Q1-K1-G6
5'-most EST
                   11244
Seq. No.
                   20996 2.R1040
Contig ID
                   LIB3094-040-Q1-K1-H1
5'-most EST
Method
                   BLASTN
NCBI GI
                   q3510336
                   73
BLAST score
                   9.0e-33
E value
                   77
Match length
                   99
% identity
                   Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:
NCBI Description
                   K18J17, complete sequence [Arabidopsis thaliana]
                   11245
Seq. No.
                   20996 3.R1040
Contig ID
                   leu701148232.hl
5'-most EST
                   11246
 Seq. No.
                   21005 1.R1040
Contig ID
 5'-most EST
                   LIB3050-005-Q1-K1-B6
                   BLASTX
Method
                   g1169017
NCBI GI
                   571
BLAST score
                   8.0e-59
E value
                   181
Match length
 % identity
                   27
                   COATOMER BETA' SUBUNIT (BETA'-COAT PROTEIN) (BETA'-COP)
 NCBI Description
                   >gi_1078361_pir__B55123 coatomer complex beta' chain -
                   yeast (Saccharomyces cerevisiae) >gi 595415 (U11237)
```

>gi 1246841 emb_CAA63359_ (X92670) see also U11237; yeast

betaprime COP [Saccharomyces cerevisiae]



coatomer subunit [Saccharomyces cerevisiae]
>gi_1322710_emb_CAA96848_ (Z72659) ORF YGL137w
[Saccharomyces cerevisiae]

Seq. No. 11247

Contig ID 21027 1.R1040

5'-most EST LIB3050-004-Q1-E1-H3

Seq. No. 11248

Contig ID 21030 1.R1040

5'-most EST LIB3050-004-Q1-E1-H6

Method BLASTX
NCBI GI g3851214
BLAST score 145
E value 8.0e-09
Match length 103

% identity 8

NCBI Description (AJ012449) NS1-binding protein [Homo sapiens]

Seq. No. 11249

Contig ID 21032 1.R1040 5'-most EST gsv701044272.h1

Seq. No. 11250

Contig ID 21032_2.R1040 5'-most EST hrw701063207.h1

Seq. No. 11251

Contig ID 21036 1.R1040

5'-most EST jC-gmle01810092g02d1

Method BLASTX
NCBI GI g4185599
BLAST score 216
E value 2.0e-17
Match length 137
% identity 37

NCBI Description (AB010708) Anthocyanin 5-aromatic acyltransferase [Gentiana

triflora]

Seq. No. 11252

Contig ID 21044 1.R1040

5'-most EST LIB3050-004-Q1-E1-H2

Method BLASTX
NCBI GI 94056438
BLAST score 247
E value 7.0e-21
Match length 120
% identity 41

NCBI Description (AC005990) F508.11 [Arabidopsis thaliana]

Seq. No. 11253

Contig ID 21050_1.R1040

5'-most EST jC-gmf102220103a11d1

Seq. No. 11254

Contig ID 21051_1.R1040

5'-most EST LIB3139-094-P1-N1-E1



```
BLASTX
Method
                  g1066447
NCBI GI
BLAST score
                  177
E value
                  3.0e-12
Match length
                  182
% identity
                  31
                  (Z67980) Chalcone isomerase [Callistephus chinensis]
NCBI Description
Seq. No.
                  11255
Contig ID
                  21051 2.R1040
                  uC-gmrominsoy212e01b1
5'-most EST
Seq. No.
                  11256
                  21057 1.R1040
Contig ID
5'-most EST
                  leu701149634.h1
Method
                  BLASTX
                  g1168940
NCBI GI
BLAST score
                  483
E value
                  3.0e-50
Match length
                  146
                  70
% identity
                  CHORISMATE MUTASE PRECURSOR (CM-1) >gi_629509 pir__S38958
NCBI Description
                  chorismate mutase precursor - Arabidopsis thaliana
                  >gi 429153 emb CAA81286 (Z26519) chorismate mutase
                  precursor [Arabidopsis thaliana]
Seq. No.
                   11257
Contig ID
                   21071 1.R1040
5'-most EST
                   uC-gmropic066e01b1
Method
                   BLASTX
                   g3901012
NCBI GI
BLAST score
                   913
                   9.0e-99
E value
                   223
Match length
                   73
% identity
                  (AJ130885) xyloglucan endotransglycosylase 1 [Fagus
NCBI Description
                   sylvatica]
                   11258
Seq. No.
                   21072 1.R1040
Contig ID
5'-most EST
                   g5753493
Seq. No.
                   11259
                   21076 2.R1040
Contig ID
                   pmv700889252.h1
5'-most EST
Seq. No.
                   11260
                   21084 1.R1040
Contig ID
```

 $ssr70\overline{0}556809.h1$ 5'-most EST

BLASTX Method NCBI GI g3757520 BLAST score 211 3.0e-16 E value Match length 169 33 % identity

NCBI Description (AC005167) unknown protein [Arabidopsis thaliana]



```
Seq. No.
                  11261
                  21102 1.R1040
Contig ID
                  LIB3139-002-Q1-N1-C12
5'-most EST
Seq. No.
                  11262
                  21102 2.R1040
Contig ID
                  jC-qmst02400037c03a1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q2827528
BLAST score
                  178
E value
                  6.0e-13
                  105
Match length
% identity
NCBI Description (AL021633) predicted protein [Arabidopsis thaliana]
Seq. No.
                  11263
                  21103 1.R1040
Contig ID
5'-most EST
                  bth700847330.h1
                  BLASTX
Method
                  q4539456
NCBI GI
BLAST score
                  710
E value
                  2.0e-79
Match length
                  215
                   62
% identity
NCBI Description (AL049500) osmotin precursor [Arabidopsis thaliana]
Seq. No.
                   11264
Contig ID
                   21103 2.R1040
5'-most EST
                   jC-qmro02910008h01a1
Method
                  BLASTX
NCBI GI
                   q129320
BLAST score
                   235
E value
                   9.0e-20
                   55
Match length
% identity
                   71
NCBI Description P21 PROTEIN >gi_99943_pir__A33176 P21 protein - soybean
                   11265
Seq. No.
                   21104 1.R1040
Contig ID
                   uC-gmropic058b10b1
5'-most EST
Method
                   BLASTN
                   g3821780
NCBI GI
BLAST score
                   36
                   1.0e-10
E value
Match length
                   36
                   100
% identity
NCBI Description Xenopus laevis cDNA clone 27A6-1
                   11266
Seq. No.
```

Contig ID 21104 2.R1040

5'-most EST uC-gmrominsoy257a10b1

Seq. No. 11267

Contig ID 21108_1.R1040 5'-most EST gsv701054978.h1

Method BLASTX NCBI GI g3874563

1949



BLAST score 589 E value 1.0e-60 Match length 253 % identity 49

NCBI Description

(Z81042) similar to Yeast hypothetical protein YEY6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk206h5.5 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA ... >gi_3924825_emb_CAB05549_ (Z83113) similar to Yeast hypothetical protein YEY6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA

Seq. No. 11268

Contig ID 21114 1.R1040

5'-most EST LIB3107-009-Q1-K1-F9

Seq. No. 11269

Contig ID 21118_1.R1040 5'-most EST leu701144831.h1

Seq. No. 11270

Contig ID 21122 1.R1040

5'-most EST LIB3109-036-Q1-K1-E1

Seq. No. 11271

Contig ID 21130 1.R1040

5'-most EST uC-gmrominsoy211e11b1

Seq. No. 11272

Contig ID 21136_1.R1040 5'-most EST eep700868362.h1

Method BLASTX NCBI GI g3080418

BLAST score 212 E value 1.0e-16 Match length 68

Match length 68 % identity 53

NCBI Description (AL022604) putative protein [Arabidopsis thaliana]

Seq. No. 11273

Contig ID 21139 1.R1040

5'-most EST uC-gmflminsoy024f01b1

Seq. No. 11274

Contig ID 21139_2.R1040 5'-most EST kl1701203774.h2

Seq. No. 11275

Contig ID 21141_1.R1040

5'-most EST uC-gmronoir048f04b1

Method BLASTN
NCBI GI g4566613
BLAST score 153
E value 3.0e-80
Match length 401



% identity Populus alba x Populus tremula actin depolymerizing factor NCBI Description mRNA, partial cds Seq. No. 11276 21141 3.R1040 Contig ID LIB3107-039-Q1-K1-C6 5'-most EST BLASTN Method g4566613 NCBI GI BLAST score 86 2.0e-40 E value Match length 270 83 % identity Populus alba x Populus tremula actin depolymerizing factor NCBI Description mRNA, partial cds 11277 Seq. No. Contig ID 21144 1.R1040 LIB3170-080-Q1-K1-D5 5'-most EST BLASTX Method NCBI GI q3786008 BLAST score 363 E value 3.0e-34 201 Match length 45 % identity (AC005499) unknown protein [Arabidopsis thaliana] NCBI Description 11278 Seq. No. Contig ID 21149 1.R1040 LIB3050-003-Q1-E1-E7 5'-most EST Seq. No. 11279 21153 1.R1040 Contig ID fjg700968328.h1 5'-most EST BLASTX Method g2245087 NCBI GI BLAST score 547 5.0e-57 E value 268 Match length 46 % identity NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana] 11280 Seq. No. 21153 2.R1040 Contig ID 5'-most EST LIB3051-026-Q1-K1-H1 BLASTX Method g2245087 NCBI GI

Method BLASTX
NCBI GI g2245087
BLAST score 176
E value 2.0e-12
Match length 42
% identity 69

NCBI Description (Z97343) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11281

Contig ID 21154_1.R1040

5'-most EST LIB3050-003-Q1-E1-D3

Method BLASTX

```
g3176708
NCBI GI
BLAST score
                  149
                  6.0e-14
E value
                  90
Match length
% identity
                  (AC002392) putative proline-rich protein APG [Arabidopsis
NCBI Description
                  thaliana]
                  11282
Seq. No.
                  21162 1.R1040
Contig ID
5'-most EST
                  k11701206442.h1
Method
                  BLASTX
                  g2982301
NCBI GI
                  285
BLAST score
                   2.0e-25
E value
Match length
                   108
% identity
NCBI Description (AF051235) YGL010w-like protein [Picea mariana]
                   11283
Seq. No.
Contig ID
                   21164 1.R1040
                   LIB3050-003-Q1-E1-E2
5'-most EST
                   BLASTX
Method
                   g2252824
NCBI GI
BLAST score
                   600
                   2.0e-62
E value
Match length
                   144
                   76
% identity
                  (AF013293) No definition line found [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   11284
                   21165 1.R1040
Contig ID
5'-most EST
                   q5678094
Method
                   BLASTN
                   g3582780
NCBI GI
                   152
BLAST score
                   1.0e-79
E value
                   440
Match length
                   84
% identity
                   Nicotiana tabacum clone 354 peroxisomal targeting sequence
NCBI Description
                   1 receptor (PEX5) mRNA, partial cds
                   11285
Seq. No.
                   21165 2.R1040
Contig ID
                   pmv700891870.h1
5'-most EST
                   BLASTX
Method
                   g3582779
NCBI GI
                   260
BLAST score
                   2.0e-22
E value
Match length
                   51
                   96
 % identity
                   (AF056282) peroxisomal targeting sequence 1 receptor
NCBI Description
                   [Nicotiana tabacum]
```

Contig ID 21165_3.R1040 5'-most EST jex700904785.h1



```
BLASTX
Method
                  q3582779
NCBI GI
                  216
BLAST score
                  8.0e-18
E value
Match length
                  42
% identity
                  95
                  (AF056282) peroxisomal targeting sequence 1 receptor
NCBI Description
                   [Nicotiana tabacum]
                  11287
Seq. No.
                  21167 1.R1040
Contig ID
                  jC-qmst02400003d04d1
5'-most EST
                  11288
Seq. No.
                  21169 1.R1040
Contig ID
                  bth700846963.h1
5'-most EST
                   11289
Seq. No.
                   21171 1.R1040
Contig ID
                   jC-gmro02910068d03a1
5'-most EST
                   BLASTX
Method
                   g3395439
NCBI GI
BLAST score
                   330
                   2.0e-30
E value
                   254
Match length
                   35
% identity
NCBI Description (AC004683) hypothetical protein [Arabidopsis thaliana]
                   11290
Seq. No.
Contig ID
                   21171 2.R1040
                   fua701042933.h1
5'-most EST
Seq. No.
                   11291
                   21171 3.R1040
Contig ID
5'-most EST
                   txt700736855.h1
                   11292
Seq. No.
Contig ID
                   21174 1.R1040
5'-most EST
                   LIB3094-006-Q1-K1-G10
                   11293
Seq. No.
Contig ID
                   21174 2.R1040
5'-most EST
                   jC-gmro02910005a07a1
                   11294
Seq. No.
                   21174 3.R1040
Contig ID
5'-most EST
                   leu701154556.hl
Seq. No.
                   11295
Contig ID
                   21175 1.R1040
                   zpv700757458.h1
5'-most EST
Method
                   BLASTX
```

NCBI GI g2194120
BLAST score 155
E value 7.0e-10
Match length 116
% identity 39



NCBI Description (AC002062) No definition line found [Arabidopsis thaliana]

Seq. No. 11296

Contig ID 21175_2.R1040 5'-most EST zhf700964993.h1

Seq. No. 11297

Contig ID 21178 1.R1040

5'-most EST LIB3093-009-Q1-K1-F4

Method BLASTX
NCBI GI g1495804
BLAST score 3624
E value 0.0e+00
Match length 926
% identity 75

NCBI Description (X96406) 13-lipoxygenase [Solanum tuberosum]

Seq. No. 11298

Contig ID 21178 2.R1040

5'-most EST uC-qmrominsoy126d02b1

Seq. No. 11299

Contig ID 21180 1.R1040

5'-most EST LIB3050-003-Q1-E1-C4

Method BLASTX
NCBI GI g2245139
BLAST score 181
E value 3.0e-13
Match length 111
% identity 45

NCBI Description (Z97344) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11300

Contig ID 21197_1.R1040

5'-most EST LIB3050-003-Q1-E1-A8

Method BLASTN
NCBI GI g516102
BLAST score 218
E value 1.0e-119
Match length 425
% identity 95

NCBI Description Soybean phytochrome B (phyB) gene exons 1-5, complete cds

Seq. No. 11301

Contig ID 21198_1.R1040 5'-most EST fC-gmro700865957a4

Method BLASTX
NCBI GI g1871192
BLAST score 879
E value 3.0e-94
Match length 554
% identity 40

NCBI Description (U90439) Cys3His zinc finger protein isolog [Arabidopsis

thaliana]

Seq. No. 11302

Contig ID 21198_2.R1040



```
LIB3050-003-Q1-E1-A9
5'-most EST
                  11303
Seq. No.
                  21198 3.R1040
Contig ID
5'-most EST
                  jC-gmro02910064b03a1
Method
                  BLASTX
                  q1871192
NCBI GI
                  441
BLAST score
                  1.0e-43
E value
                  161
Match length
% identity
NCBI Description (U90439) Cys3His zinc finger protein isolog [Arabidopsis
                  thaliana]
                  11304
Seq. No.
                  21198 5.R1040
Contig ID
                  LIB3139-006-P1-N1-D12
5'-most EST
                  11305
Seq. No.
                  21206 1.R1040
Contig ID
                  LIB3050-002-Q1-E1-G6
5'-most EST
                  11306
Seq. No.
                  21208 1.R1040
Contig ID
                   jC-gmro02800032a08a1
5'-most EST
                  BLASTX
Method
                   g2664202
NCBI GI
                   370
BLAST score
                   1.0e-34
E value
                   399
Match length
% identity
NCBI Description (AJ003217) GT2 [Arabidopsis thaliana]
                   11307
Seq. No.
                   21208 2.R1040
Contig ID
                   LIB3170-044-Q1-K1-C2
5'-most EST
                   BLASTX
Method
                   g2664204
NCBI GI
                   152
BLAST score
                   7.0e-10
E value
Match length
                   77
% identity
                   35
NCBI Description (AJ003218) GTL1 [Arabidopsis thaliana]
                   11308
Seq. No.
                   21209 1.R1040
Contig ID
                   LIB3138-125-Q1-N1-B4
5'-most EST
Method
                   BLASTX
                   g416639
NCBI GI
                   282
BLAST score
                   7.0e-25
E value
                   87
Match length
 % identity
                   68
                   INDOLE-3-ACETIC ACID INDUCED PROTEIN ARG2
 NCBI Description
                   >gi_287564_dbj_BAA03307_ (D14411) ORF [Vigna radiata]
```



```
21210 1.R1040
Contig ID
                  LIB30\overline{5}1-116-Q1-K1-G5
5'-most EST
                   BLASTX
Method
NCBI GI
                   g3176686
                   807
BLAST score
                   2.0e-86
E value
                   267
Match length
                   58
% identity
                   (AC003671) Similar to high affinity potassium transporter,
NCBI Description
                   HAK1 protein gb_U22945 from Schwanniomyces occidentalis.
                   [Arabidopsis thaliana]
                   11310
Seq. No.
                   21212_1.R1040
Contig ID
5'-most EST
                   leu701146821.h1
Seq. No.
                   11311
                   21213 1.R1040
Contig ID
                   LIB3050-002-Q1-E1-H12
5'-most EST
                   11312
Seq. No.
Contig ID
                   21216 1.R1040
5'-most EST
                   eep700870341.hl
                   BLASTX
Method
NCBI GI
                   g2467274
BLAST score
                   337
E value
                   2.0e-31
Match length
                   148
% identity
                   53
                   (Z99759) rna binding protein [Schizosaccharomyces pombe]
NCBI Description
Seq. No.
                   11313
                   21217 1.R1040
Contig ID
                   LIB3050-002-Q1-E1-H6
5'-most EST
                   11314
Seq. No.
                   21218 1.R1040
Contig ID
                   zsg701129457.h1
5'-most EST
                   11315
Seq. No.
                   21223 1.R1040
Contig ID
                   LIB3050-002-Q1-E1-F11
5'-most EST
                   11316
Seq. No.
                   21224 1.R1040
Contig ID
                   jC-gmro02910075c09a1
5'-most EST
                   BLASTX
Method
                   g3927836
NCBI GI
                   321
BLAST score
                   1.0e-37
E value
                   200
Match length
 % identity
                   47
                   (AC005727) unknown protein [Arabidopsis thaliana]
```

NCBI Description

11317

Contig ID

21224 2.R1040

5'-most EST

jC-gmro02910004g02a1



Method BLASTX
NCBI GI g3927836
BLAST score 354
E value 1.0e-43
Match length 118
% identity 80

NCBI Description (AC005727) unknown protein [Arabidopsis thaliana]

Seq. No. 11318

Contig ID 21244_1.R1040 5'-most EST uC-gmronoir016h05b1

Method BLASTX
NCBI GI g2292988
BLAST score 191
E value 1.0e-13
Match length 366
% identity 26

NCBI Description (Y11283) Inter-alpha-inhibitor H4 heavy chain [Rattus

norvegicus]

Seq. No. 11319

Contig ID 21244 3.R1040

5'-most EST jC-gmf102220057a09d1

Seq. No. 11320

Contig ID 21248_1.R1040 5'-most EST zsg701129729.h1

Seq. No. 11321

Contig ID 21252_1.R1040 5'-most EST gsv701051533.h1

Method BLASTX
NCBI GI 94218535
BLAST score 626
E value 5.0e-65
Match length 201
% identity 59

NCBI Description (AJ010829) GRAB1 protein [Triticum sp.]

Seq. No. 11322

Contig ID 21254_1.R1040 5'-most EST gsv701045606.h1

Method BLASTX
NCBI GI g3970880
BLAST score 448
E value 2.0e-44
Match length 179
% identity 49

NCBI Description (AB015724) nuclear receptor binding factor-1 [Rattus

norvegicus]

Seq. No. 11323

Contig ID 21261 1.R1040

5'-most EST uC-gmflminsoy053g04b1

Method BLASTX
NCBI GI g3522938
BLAST score 896

1957

```
2.0e-96
E value
                   401
Match length
% identity
                   (AC004411) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   11324
                   21263 1.R1040
Contig ID
                   zhf700963542.hl
5'-most EST
                   BLASTX
Method
                   q2760086
NCBI GI
BLAST score
                   425
                   3.0e-52
E value
                   178
Match length
                   19
% identity
NCBI Description (Y16046) leucine-rich repeat protein [Arabidopsis thaliana]
Seq. No.
                   11325
                   21265 1.R1040
Contig ID
                   LIB3053-008-Q1-N1-G11
5'-most EST
                   BLASTX
Method
NCBI GI
                   q1480347
BLAST score
                   1788
                   0.0e+00
E value
                   378
Match length
                   87
% identity
                   (X99419) ferrodoxin NADP oxidoreductase [Pisum sativum]
NCBI Description
Seq. No.
                   11326
                   21265 2.R1040
Contig ID
5'-most EST
                   asj700967432.hl
                   BLASTN
Method
                   q1480346
NCBI GI
BLAST score
                   57
                   3.0e-23
E value
Match length
                   129
                   86
% identity
NCBI Description P.sativum mRNA for ferrodoxin NADP oxidoreductase
                   11327
Seq. No.
                   21265 3.R1040
Contig ID
                   jC-qmst02400064a05d1
5'-most EST
Seq. No.
                   11328
                   21267 1.R1040
Contig ID
                   crh700854281.h1
5'-most EST
                   11329
Seq. No.
                   21269 1.R1040
Contig ID
                   LIB3050-002-Q1-E1-D10
5'-most EST
                   11330
Seq. No.
                   21270 1.R1040
Contig ID
                   r1r70\overline{0}901075.h1
 5'-most EST
                   BLASTX
Method
                   g1742951
NCBI GI
```

1958

902

1.0e-97

BLAST score E value



Match length 201 % identity 83

NCBI Description (Y09817) Ca2+-ATPase [Arabidopsis thaliana]

Seq. No. 11331

Contig ID 21272_1.R1040

5'-most EST LIB3139-013-P1-N1-G6

Method BLASTX
NCBI GI g3219273
BLAST score 585
E value 3.0e-60
Match length 242
% identity 48

NCBI Description (AB015316) MAP kinase kinase 5 [Arabidopsis thaliana]

Seq. No. 11332

Contig ID 21277_1.R1040

5'-most EST LIB3050-002-Q1-E1-A6

Seq. No. 11333

Contig ID 21279_1.R1040

5'-most EST LIB3050-002-Q1-E1-B1

Seq. No. 11334

Contig ID 21282_1.R1040

5'-most EST LIB3050-002-Q1-E1-B2

Method BLASTX
NCBI GI g1730796
BLAST score 347
E value 9.0e-33
Match length 146
% identity 46

NCBI Description HYPOTHETICAL 124.5 KD PROTEIN IN SKO1-RPL44A INTERGENIC

REGION >gi_2132760_pir__S60964 probable membrane protein

YNL163c - yeast (Saccharomyces cerevisiae)

>gi_1050790 emb_CAA63276_ (X92517) N1718 [Saccharomyces
cerevisiae] >gi_1302132_emb_CAA96050_ (Z71439) ORF YNL163c

[Saccharomyces cerevisiae]

Seq. No. 11335

Contig ID 21286_1.R1040

5'-most EST LIB3170-067-Q1-K1-C8

Method BLASTX
NCBI GI g4008547
BLAST score 147
E value 4.0e-09
Match length 100
% identity 33

NCBI Description (AL022072) arginine n-methyltransferase

[Schizosaccharomyces pombe]

Seq. No. 11336

Contig ID 21288_1.R1040

5'-most EST LIB3051-079-Q1-K1-D4

Seq. No. 11337

Contig ID 21296 1.R1040

1959



```
LIB3139-119-P1-N1-B6
5'-most EST
                    BLASTX
Method
                    g3641252
NCBI GI
BLAST score
                    421
E value
                    2.0e-41
Match length
                    135
                    13
% identity
                    (AF053127) leucine-rich receptor-like protein kinase [Malus
NCBI Description
                    domestica]
Seq. No.
                    11338
                    21297 1.R1040
Contig ID
                    g5126291
5'-most EST
                    BLASTX
Method
                    g122087
NCBI GI
BLAST score
                     632
E value
                     4.0e-66
                    136
Match length
                     95
% identity
                    HISTONE H3 >gi_81849_pir__S04520 histone H3 (clone pH3c-1) - alfalfa >gi_82609_pir__A26014 histone H3 - wheat >gi_19607_emb_CAA31964__(X13673) histone H3 (AA 1-136)
NCBI Description
                     [Medicago sativa] >gi_19609_emb_CAA31965_
                                                                    (X13674) histone
                     H3 (AA 1-136) [Medicago sativa] >gi_21797_emb_CAA25451
                     (X00937) H3 histone [Triticum aestivum] >gi 488565 (U09459)
                     histone H3.1 [Medicago sativa] >gi_2565419 (AF026803)
                     histone H3 [Onobrychis viciifolia]
Seq. No.
                     11339
                     21298 1.R1040
Contig ID
                     LIB3050-001-Q1-E1-F11
5'-most EST
                     11340
Seq. No.
                     21298 2.R1040
Contig ID
                     LIB3170-055-Q1-K1-A8
5'-most EST
                     11341
Seq. No.
                     21300 1.R1040
Contig ID
                     jsh701063920.h1
5'-most EST
                     11342
 Seq. No.
                     21304 1.R1040
Contig ID
                     LIB3050-001-Q1-E1-F9
 5'-most EST
                     11343
 Seq. No.
 Contig ID
                     21305 1.R1040
                     LIB3050-001-Q1-E1-G1
 5'-most EST
                     BLASTX
 Method
 NCBI GI
                     q282980
                     373
 BLAST score
 E value
                     8.0e-36
 Match length
                     99
```

72

% identity

NCBI Description

(M84913) DEL [Antirrhinum majus]

helix-loop-helix protein DEL - garden snapdragon >gi_166428



Contig ID 21306_1.R1040

5'-most EST LIB3051-001-Q1-E1-G11

Method BLASTX
NCBI GI g1169782
BLAST score 615
E value 1.0e-129
Match length 290
% identity 80

NCBI Description FUSCA PROTEIN FUS6 >gi_432446 (L26498) FUS6 [Arabidopsis

thaliana]

Seq. No. 11345

Contig ID 21310_1.R1040 5'-most EST awf700840781.h1

Method BLASTX
NCBI GI g2388562
BLAST score 166
E value 2.0e-11
Match length 73
% identity 47

NCBI Description (AC000098) EST gb_ATTS0295 comes from this gene.

[Arabidopsis thaliana]

Seq. No. 11346

Contig ID 21314_1.R1040 5'-most EST awf700842001.h1

Method BLASTX
NCBI GI g2392056
BLAST score 738
E value 5.0e-78
Match length 357
% identity 44

NCBI Description Chain A, Class 3 Aldehyde Dehydrogenase Complex With Nicotinamide-Adenine-Dinucleotide >gi_2392057_pdb_1AD3_B

Chain B, Class 3 Aldehyde Dehydrogenase Complex With

Nicotinamide-Adenine-Dinucleotide

Seq. No. 11347

Contig ID 21314 2.R1040

5'-most EST uC-gmrominsoy062b01b1

Method BLASTX
NCBI GI g4455169
BLAST score 166
E value 1.0e-11
Match length 66
% identity 48

NCBI Description (AL035521) putative aldehyde dehydrogenase [Arabidopsis

thaliana]

Seq. No. 11348

Contig ID 21316_1.R1040

5'-most EST LIB3050-001-Q1-E1-E11

Seq. No. 11349

Contig ID 21319_1.R1040

5'-most EST LIB3170-056-Q1-K1-D5



Contig ID 21319 2.R1040

5'-most EST LIB3051-001-Q1-E1-E5

Seq. No. 11351

Contig ID 21323_1.R1040

5'-most EST LIB3107-014-Q1-K1-C1

Seq. No. 11352

Contig ID 21325 1.R1040

5'-most EST LIB3051-001-Q1-E1-C12

Seq. No. 11353

Contig ID 21327 1.R1040

5'-most EST LIB $30\overline{5}0-001-Q1-E1-C6$

Method BLASTX
NCBI GI 94490702
BLAST score 143
E value 2.0e-13
Match length 109

Match length 109 % identity 52

NCBI Description (AL035680) hypothetical protein [Arabidopsis thaliana]

Seq. No. 11354

Contig ID 21333 1.R1040

5'-most EST jC-gmro02910040g04a1

Seq. No. 11355

Contig ID 21338 1.R1040

5'-most EST LIB3051-086-Q1-K1-E10

Seq. No. 11356

Contig ID 21341 1.R1040 5'-most EST lus701015777.h1

Seq. No. 11357

Contig ID 21342_1.R1040 5'-most EST bth700846441.h1

Method BLASTX
NCBI GI g135535
BLAST score 2385
E value 0.0e+00
Match length 545
% identity 88

NCBI Description T-COMPLEX PROTEIN 1, ALPHA SUBUNIT (TCP-1-ALPHA)

(CCT-ALPHA) >gi_322602_pir__JN0448 t-complex polypeptide Tcp-1 - Arabidopsis thaliana >gi_217871_dbj_BAA01955_(D11351) t-complex polypeptide 1 homologue [Arabidopsis

thaliana] >gi_2326265_dbj_BAA21772_ (D11352) CCT

alpha/TCP-1 [Arabidopsis thaliana]

Seq. No. 11358

Contig ID 21345_1.R1040 5'-most EST kl1701206501.h1

Method BLASTX
NCBI GI g116923
BLAST score 931



E value 1.0e-100 Match length 332 % identity 59

NCBI Description COATOMER BETA SUBUNIT (BETA-COAT PROTEIN) (BETA-COP)

>gi_111414_pir__S13520 beta-COP protein - rat
>gi_55819_emb_CAA40505_ (X57228) beta COP [Rattus

norvegicus]

Seq. No. 11359

Contig ID 21349_1.R1040

5'-most EST LIB3051-012-Q1-E1-H1

Method BLASTX
NCBI GI g2134962
BLAST score 321
E value 2.0e-29
Match length 119
% identity 53

NCBI Description cyclophilin-like protein CyP-60 - human >gi_1199598

(U37219) cyclophilin-like protein CyP-60 [Homo sapiens] >gi_1587642_prf__2207180A cyclophilin:ISOTYPE=CyP-60 [Homo

sapiens]

Seq. No. 11360

Contig ID 21350_1.R1040

5'-most EST LIB3050-001-Q1-E1-A9

Seq. No. 11361

Contig ID 21351_1.R1040

5'-most EST LIB3050-001-Q1-E1-B2

Seq. No. 11362

Contig ID 21354 1.R1040

5'-most EST LIB3050-001-Q1-E1-C1

Seq. No. 11363

Contig ID 21355_1.R1040

5'-most EST LIB3050-001-Q1-E1-C11

Seq. No. 11364

Contig ID 21358_1.R1040 5'-most EST cle700967790.h1

Seq. No. 11365

Contig ID 21358_2.R1040

5'-most EST LIB3139-028-P1-N1-D3

Seq. No. 11366

Contig ID 21359_1.R1040 5'-most EST bth700843780.h1

Method BLASTX
NCBI GI g3928869
BLAST score 163
E value 4.0e-11
Match length 102
% identity 38

NCBI Description (AF093420) Hsp70 binding protein HspBP1 [Homo sapiens]



Seq. No. 11367 21362 1.R1040 Contig ID LIB3050-001-Q1-E1-E2 5'-most EST Seq. No. 11368 Contig ID 21364 1.R1040 LIB3051-008-Q1-E1-F12 5'-most EST 11369 Seq. No. 21364 2.R1040 Contig ID LIB3093-032-Q1-K1-E10 5'-most EST 11370 Seq. No. 21365 1.R1040 Contig ID zpv700763416.h1 5'-most EST Method BLASTN NCBI GI g1065907 BLAST score 387 E value 0.0e + 00959 Match length 90 % identity Pisum sativum chloroplast processing enzyme mRNA, nuclear NCBI Description gene encoding chloroplast protein, complete cds 11371 Seq. No. 21370 1.R1040 Contig ID 5'-most EST gsv701048554.h1 BLASTX Method q1350548 NCBI GI BLAST score 176 2.0e-12 E value 60 Match length % identity 58 (L47609) heat shock-like protein [Picea glauca] NCBI Description 11372 Seq. No. Contig ID 21371 1.R1040 LIB3050-001-Q1-E1-G5 5'-most EST Method BLASTX q4336436 NCBI GI BLAST score 373 1.0e-35 E value Match length 114 65 % identity (AF092432) protein phosphatase type 2C [Lotus japonicus] NCBI Description 11373 Seq. No. 21372 1.R1040 Contig ID LIB3050-001-Q1-E1-H11 5'-most EST

Method BLASTX
NCBI GI g421955
BLAST score 155
E value 2.0e-16
Match length 106
% identity 54

NCBI Description hypothetical protein 4 - potato transposon Tst1

>gi 21434 emb CAA36616 (X52387) ORF4 [Solanum tuberosum]

```
11374
Seq. No.
                   21373 1.R1040
Contig ID
5'-most EST
                  LIB3050-001-Q1-E1-H12
Method
                   BLASTX
NCBI GI
                   q2980806
BLAST score
                   588
E value
                   7.0e-61
                   176
Match length
                   48
% identity
NCBI Description (AL022197) putative protein [Arabidopsis thaliana]
                   11375
Seq. No.
Contig ID
                   21375 1.R1040
                   jC-gmf102220083a02a1
5'-most EST
Method
                   BLASTX
NCBI GI
                   q3643598
BLAST score
                   325
                   2.0e-29
E value
                   89
Match length
                   62
% identity
                   (AC005395) putative poly(A) polymerase [Arabidopsis
NCBI Description
                   thaliana]
                   11376
Seq. No.
                   21376 1.R1040
Contig ID
5'-most EST
                   jC-gmro02800030a01d1
                   11377
Seq. No.
Contig ID
                   21383 1.R1040
                   LIB3107-029-Q1-K1-F7
5'-most EST
                   BLASTX
Method
                   q2980767
NCBI GI
BLAST score
                   533
                   3.0e-54
E value
Match length
                   170
% identity
                   63
                   (AL022198) putative protein [Arabidopsis thaliana]
NCBI Description
                   11378
Seq. No.
                   21386 1.R1040
Contig ID
                   q4295843
5'-most EST
                   11379
Seq. No.
                   21386 2.R1040
Contig ID
                   95606\overline{5}41
5'-most EST
                   BLASTX
Method
                   g1174199
NCBI GI
                   216
BLAST score
                   4.0e-17
E value
                   130
Match length
% identity
                   45
                   (U44760) S25-PR6 [Nicotiana tabacum]
NCBI Description
                   11380
Seq. No.
```

1965

21404 1.R1040

zsg701118990.hl

Contig ID

5'-most EST

Method

NCBI GI

E value Match length

BLAST score



11381 Seq. No. 21410 1.R1040 Contig ID txt700734121.hl 5'-most EST BLASTX Method NCBI GI a4580389 454 BLAST score 6.0e - 45E value 228 Match length 43 % identity NCBI Description (AC007171) unknown protein [Arabidopsis thaliana] 11382 Seq. No. 21410 2.R1040 Contig ID 5'-most EST bth700849572.h1 BLASTX Method g4580389 NCBI GI BLAST score 316 9.0e-29 E value 159 Match length 41 % identity (AC007171) unknown protein [Arabidopsis thaliana] NCBI Description 11383 Seq. No. Contig ID 21414 1.R1040 LIB3106-085-Q1-K1-B6 5'-most EST BLASTX Method q4538973 NCBI GI 553 BLAST score E value 1.0e-56 218 Match length % identity 51 (ALO49487) hypothetical protein [Arabidopsis thaliana] NCBI Description 11384 Seq. No. 21419 1.R1040 Contig ID LIB3050-024-Q1-K1-D11 5'-most EST Method BLASTX q1532167 NCBI GI 150 BLAST score 1.0e-09 E value Match length 65 % identity 38 (U63815) localized according to blastn similarity to EST NCBI Description sequences; therefore, the coding span corresponds only to an area of similarity since the initation codon and stop codon could not be precisely determined [Arabidopsis thaliana] 11385 Seq. No. 21423 1.R1040 Contig ID 5'-most EST $g5175\overline{4}63$ BLASTX

1966

g2191141

1.0e-85

691

266

```
% identity
                  (AF007269) A IG002N01.21 gene product [Arabidopsis
NCBI Description
                  thaliana]
                  11386
Seq. No.
Contig ID
                  21425 1.R1040
                  hrw701062017.h1
5'-most EST
                  BLASTX
Method
                  q3386565
NCBI GI
                  1028
BLAST score
                   1.0e-112
E value
                   307
Match length
% identity
                   (AF079588) 1-aminocyclopropane-1-carboxylate oxidase
NCBI Description
                   [Sorghum bicolor]
                   11387
Seq. No.
Contig ID
                   21425 2.R1040
                   jC-qmst02400007d12d1
5'-most EST
                   BLASTX
Method
                   a398994
NCBI GI
BLAST score
                   302
                   3.0e-27
E value
                   89
Match length
% identity
                   63
                   1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE (ACC OXIDASE)
NCBI Description
                   (ETHYLENE-FORMING ENZYME) (EFE) >gi_166313 (M97961) tomato
                   and apple ACC oxidase homologue [Actinidia deliciosa]
Seq. No.
                   11388
                   21429 1.R1040
Contig ID
                   LIB30\overline{5}0-028-Q1-E1-B2
5'-most EST
Seq. No.
                   11389
                   21437 1.R1040
Contig ID
                   LIB3051-010-Q1-E1-G8
5'-most EST
Method
                   BLASTX
                   g3834302
NCBI GI
                   2334
BLAST score
                   0.0e + 00
E value
Match length
                   559
% identity
                   83
                   (ACO05679) Similar to gb D45384 vacuolar H+-pyrophosphatase
NCBI Description
                   from Oryza sativa. ESTs gb_F14272 and gb_F14273 come from
                   this gene. [Arabidopsis thaliana]
                   11390
Seq. No.
                   21437 2.R1040
Contig ID
                   ssr700558351.hl
5'-most EST
                   BLASTX
Method
```

Method BLASTX
NCBI GI g3355717
BLAST score 551
E value 2.0e-73
Match length 184
% identity 76

NCBI Description (Y13053) seryl-tRNA synthetase [Zea mays]



```
11391
Seq. No.
                  21439 1.R1040
Contig ID
                  taw700655923.hl
5'-most EST
                  BLASTX
Method
                  q3334223
NCBI GI
                  435
BLAST score
                  5.0e-43
E value
                  94
Match length
                  86
% identity
                  4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (4HPPD) (HPD)
NCBI Description
                  >gi_2145039 (AF000228) p-hydroxyphenylpyruvate dioxygenase
                  [Arabidopsis thaliana] >gi_2392518 (U89267)
                  p-hydroxyphenylpyruvate dioxygenase [Arabidopsis thaliana]
                  >gi 3098559 (AF047834) 4-hydroxyphenylpyruvate dioxygenase
                  [Arabidopsis thaliana]
                  11392
Seq. No.
                  21440 1.R1040
Contig ID
                  zsq701123794.h1
5'-most EST
                  BLASTX
Method
                  g3334219
NCBI GI
                  939
BLAST score
                  1.0e-101
E value
                  306
Match length
% identity
                  62
                  4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (4HPPD) (HPD)
NCBI Description
                  >qi 2231615 (U87257) 4-hydroxyphenylpyruvate dioxygenase
                  [Daucus carota]
                  11393
Seq. No.
                  21447 1.R1040
Contig ID
                  LIB3093-044-Q1-K1-F6
5'-most EST
                  BLASTX
Method
                  g3548818
NCBI GI
                  1391
BLAST score
                  1.0e-154
E value
                  342
Match length
                  79
% identity
NCBI Description (AC005313) unknown protein [Arabidopsis thaliana]
                  11394
Seq. No.
                  21447 2.R1040
Contig ID
                  LIB3107-034-Q1-K1-E7
5'-most EST
                  BLASTX
Method
                   g3548818
NCBI GI
BLAST score
                   332
                   5.0e-31
E value
                   94
Match length
% identity
                  71
```

Contig ID 21448_1.R1040 5'-most EST zzp700834891.h1

Method BLASTX
NCBI GI g4457221
BLAST score 298

NCBI Description (AC005313) unknown protein [Arabidopsis thaliana]



```
1.0e-26
E value
Match length
                  130
% identity
                  53
                  (AF127797) putative bZIP DNA-binding protein [Capsicum
NCBI Description
                  chinense]
                  11396
Seq. No.
                  21449 1.R1040
Contig ID
5'-most EST
                  q5057953
Method
                  BLASTX
                  g100226
NCBI GI
BLAST score
                  635
                  5.0e-66
E value
Match length
                  212
% identity
                  57
                  hypothetical protein - tomato >gi 19275 emb CAA78112
NCBI Description
                   (Z12127) protein of unknown function [Lycopersicon
                   esculentum] >gi 445619 prf 1909366A Leu zipper protein
                   [Lycopersicon esculentum]
Seq. No.
                  11397
Contig ID
                  21450 1.R1040
                  LIB3051-041-Q1-K1-D6
5'-most EST
Method
                  BLASTX
                  g3790554
NCBI GI
BLAST score
                  208
E value
                  2.0e-16
Match length
                  119
% identity
                  (AF078683) RING-H2 finger protein RHA1a [Arabidopsis
NCBI Description
                  thaliana]
                  11398
Seq. No.
                  21451 1.R1040
Contig ID
5'-most EST
                  LIB3139-088-P1-N1-G10
Method
                  BLASTX
                  g3451072
NCBI GI
BLAST score
                   616
                   9.0e-64
E value
Match length
                   158
% identity
                  (AL031326) putative protein [Arabidopsis thaliana]
NCBI Description
                   11399
Seq. No.
                   21451 3.R1040
Contig ID
5'-most EST
                  uC-gmropic070g07b1
                   BLASTX
Method
NCBI GI
                  g3451072
BLAST score
                   253
                   1.0e-21
E value
Match length
                   73
% identity
                   63
                  (AL031326) putative protein [Arabidopsis thaliana]
NCBI Description
                   11400
Seq. No.
```

Contig ID 21451 4.R1040

5'-most EST uC-gmropic022a05b1



```
Method
                   BLASTX
NCBI GI
                   q3451072
BLAST score
                   350
E value
                   8.0e-37
Match length
                   99
% identity
                   75
NCBI Description
                   (AL031326) putative protein [Arabidopsis thaliana]
Seq. No.
                   11401
Contig ID
                   21451 7.R1040
5'-most EST
                  seb70\overline{0}651120.h1
Seq. No.
                   11402
Contig ID
                   21453 1.R1040
5'-most EST
                  rlr700897774.h1
Method
                  BLASTX
NCBI GI
                   g1872521
BLAST score
                   585
E value
                   4.0e-60
Match length
                   178
% identity
NCBI Description
                   (U87833) zinc-finger protein Lsd1 [Arabidopsis thaliana]
                   >gi 1872523 (U87834) zinc-finger protein Lsd1 [Arabidopsis
                   thaliana]
Seq. No.
                   11403
                   21453 2.R1040
Contig ID
5'-most EST
                  LIB3109-001-Q1-K2-B12
Method
                  BLASTX
NCBI GI
                  g1872521
BLAST score
                   247
                   6.0e-21
E value
Match length
                  84
                   30
% identity
NCBI Description
                   (U87833) zinc-finger protein Lsd1 [Arabidopsis thaliana]
                  >gi 1872523 (U87834) zinc-finger protein Lsd1 [Arabidopsis
                  thaliana]
Seq. No.
                  11404
Contig ID
                  21453 3.R1040
5'-most EST
                  LIB3051-019-Q1-E1-H6
Method
                  BLASTX
NCBI GI
                  g1872521
BLAST score
                  500
E value
                  1.0e-62
Match length
                  178
% identity
                  50
NCBI Description
                   (U87833) zinc-finger protein Lsd1 [Arabidopsis thaliana]
                  >gi_1872523 (U87834) zinc-finger protein Lsd1 [Arabidopsis
                  thaliana]
Seq. No.
                  11405
```

Contig ID

21453 4.R1040

5'-most EST

LIB3109-043-Q1-K1-D3

Seq. No.

11406

Contig ID

21454_1.R1040



bth700849155.h1 5'-most EST Method BLASTX NCBI GI q2129575 BLAST score 557 E value 3.0e-57 Match length 142 77 % identity NCBI Description DNA repair protein homolog XPBara - Arabidopsis thaliana Seq. No. 11407 21457 1.R1040 Contig ID 5'-most EST wrg700789425.h2 BLASTX Method NCBI GI g2914706 BLAST score 648

7.0e-68 E value Match length 167 71 % identity

NCBI Description (AC003974) putative homeobox protein [Arabidopsis thaliana]

Seq. No. 11408

Contig ID 21457 2.R1040 5'-most EST leu701147230.hl

BLASTX Method g2914706 NCBI GI BLAST score 592 E value 3.0e-61 Match length 152 % identity 72

NCBI Description (AC003974) putative homeobox protein [Arabidopsis thaliana]

11409 Seq. No.

21460 1.R1040 Contig ID

5'-most EST jC-gmf102220054g09a1

Method BLASTX NCBI GI q1076660 BLAST score 1141 1.0e-125 E value 333 Match length % identity 72

NCBI Description D13F(MYBST1) protein - potato >gi 786426 bbs 159122 (S74753) MybSt1=Myb-related transcriptional activator {DNA-binding domain repeats} [Solanum tuberosum=potatoes,

leaf, Peptide, 342 aa] [Solanum tuberosum]

Seq. No. 11410

21460 3.R1040 Contig ID 5'-most EST ek1700968295.h1

Method BLASTX NCBI GI g1076660 BLAST score 154 E value 2.0e-10 28 Match length 96 % identity

NCBI Description D13F(MYBST1) protein - potato >gi 786426 bbs 159122 (S74753) MybStl=Myb-related transcriptional activator

{DNA-binding domain repeats} [Solanum tuberosum=potatoes,



leaf, Peptide, 342 aa] [Solanum tuberosum]

11411 Seq. No. Contig ID 21464 1.R1040 uxk700668564.h1 5'-most EST Method BLASTX q3152606 NCBI GI 674 BLAST score 1.0e-70 E value 169 Match length 72 % identity (AC004482) putative ring zinc finger protein [Arabidopsis NCBI Description thaliana] 11412 Seq. No. 21464 3.R1040 Contig ID LIB3051-002-Q1-E1-E5 5'-most EST BLASTX Method q3152606 NCBI GI 291 BLAST score 4.0e-26 E value Match length 54 % identity (AC004482) putative ring zinc finger protein [Arabidopsis NCBI Description thaliana] 11413 Seq. No. 21465 1.R1040 Contig ID eep700868572.h1 5'-most EST Method BLASTX g1871192 NCBI GI BLAST score 213 6.0e-17 E value 92 Match length 48 % identity (U90439) Cys3His zinc finger protein isolog [Arabidopsis NCBI Description thaliana] 11414 Seq. No. 21465 2.R1040 Contig ID uC-gmflminsoy011d06b1 5'-most EST BLASTX Method g1871192 NCBI GI 227 BLAST score 1.0e-18 E value 94 Match length 49 % identity (U90439) Cys3His zinc finger protein isolog [Arabidopsis NCBI Description thaliana] 11415 Seq. No. 21466_1.R1040 Contig ID

5'-most EST xpa700793076.h1
Method BLASTX
NCBI GI g1519680

BLAST score 193 E value 3.0e-14

1972



```
Match length
                  182
                  31
% identity
NCBI Description
                  (U67953) contains similarity to C3HC4-class zinc finger
                   (PS:PS00518) [Caenorhabditis elegans]
Seq. No.
                  11416
                  21469 1.R1040
Contig ID
5'-most EST
                  LIB3051-026-Q1-K1-B7
Method
                  BLASTX
NCBI GI
                  q3152598
                  389
BLAST score
E value
                  1.0e-37
Match length
                  113
                  59
% identity
                  (AC002986) Contains similarity to C2-HC type zinc finger
NCBI Description
                  protein C.e-MyT1 gb_U67079 from C. elegans and to
                  hypersensitivity-related gene 201 isolog T28M21.14 from A.
                  thaliana BAC gb_AF002109. [Arabidopsis thaliana]
                  11417
Seq. No.
                  21472 1.R1040
Contig ID
5'-most EST
                  LIB3051-059-Q1-K2-E10
                  BLASTX
Method
NCBI GI
                  g2130078
BLAST score
                  251
E value
                  2.0e-21
Match length
                  136
% identity
                  46
NCBI Description MADS-box protein 3 - rice >gi 886405 (L37528) MADS box
                  protein [Oryza sativa]
                  11418
Seq. No.
                  21472 2.R1040
Contig ID
5'-most EST
                  LIB3170-050-Q1-J1-C8
Seq. No.
                  11419
Contig ID
                  21475_1.R1040
                  jC-gmro02910060a03a1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2865394
BLAST score
                  185
                  2.0e-13
E value
Match length
                  118
% identity
                  50
NCBI Description (AF036949) basic leucine zipper protein [Zea mays]
Seq. No.
                  11420
Contig ID
                  21476 1.R1040
5'-most EST
                  awf700836445.h1
                  BLASTX
Method
```

Method BLASTX
NCBI GI g2062176
BLAST score 859
E value 4.0e-92
Match length 356
% identity 55

NCBI Description (AC001645) Myb-related transcription activator (MybSt1) isolog [Arabidopsis thaliana]

5'-most EST

Method

NCBI GI

```
Seq. No.
                   11421
                   21476 2.R1040
Contig ID
5'-most EST
                  LIB3170-024-Q1-K1-G1
Method
                  BLASTX
NCBI GI
                   g2062176
BLAST score
                  314
E value
                   1.0e-28
Match length
                  117
% identity
                   56
NCBI Description (AC001645) Myb-related transcription activator (MybSt1)
                   isolog [Arabidopsis thaliana]
                   11422
Seq. No.
                  21476_3.R1040
Contig ID
5'-most EST
                  g5752602
Seq. No.
                   11423
Contig ID
                   21478 1.R1040
                  LIB3170-052-Q1-K1-C9
5'-most EST
Method
                  BLASTX
NCBI GI
                   g1945281
BLAST score
                   247
                   6.0e-21
E value
                   45
Match length
                   91
% identity
NCBI Description (Y11352) myb factor [Oryza sativa]
Seq. No.
                   11424
                   21478 2.R1040
Contig ID
                  LIB3051-084-Q1-K1-A12
5'-most EST
Method
                  BLASTX
                   q3941528
NCBI GI
BLAST score
                   458
E value
                   3.0e-45
                   220
Match length
% identity
                   68
NCBI Description (AF062918) putative transcription factor [Arabidopsis
                   thaliana]
                   11425
Seq. No.
                   21478 3.R1040
Contig ID
5'-most EST
                   uxk700673133.hl
Method
                   BLASTX
                   g3941528
NCBI GI
BLAST score
                   433
E value
                   1.0e-42
Match length
                   203
                   46
% identity
                  (AF062918) putative transcription factor [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   11426
Contig ID
                   21482 1.R1040
```

LIB3106-010-Q1-K1-B6

BLASTX

q2826884

Contig ID

11431

21492 1.R1040



```
BLAST score
                  393
                  5.0e-38
E value
Match length
                  169
% identity
                  54
NCBI Description
                   (AJ223635) transcription factor IIA large subunit
                  [Arabidopsis thaliana]
Seq. No.
                  11427
                  21485 1.R1040
Contig ID
5'-most EST
                  awf700839210.hl
Method
                  BLASTX
NCBI GI
                  g3600059
BLAST score
                  807
E value
                  2.0e-86
Match length
                  189
% identity
                  76
NCBI Description
                  (AF080120) contains similarity to WB domains, G-beta
                  repeats (Pfam: G-beta.hmm, score: 14.83 and 23.03)
                  [Arabidopsis thaliana]
Seq. No.
                  11428
Contiq ID
                  21488 1.R1040
5'-most EST
                  gsv701046512.h1
Method
                  BLASTX
NCBI GI
                  g4324495
BLAST score
                  414
E value
                  2.0e-40
Match length
                  140
% identity
NCBI Description
                  (AF105221) glutamyl-tRNA reductase precursor [Glycine max]
Seq. No.
                  11429
                  21490 1.R1040
Contig ID
5'-most EST
                  dpv701099817.h1
Method
                  BLASTX
NCBI GI
                  g4510377
BLAST score
                  309
E value
                  3.0e-28
Match length
                  140
% identity
NCBI Description
                  (AC007017) putative RNA helicase A [Arabidopsis thaliana]
Seq. No.
                  11430
Contig ID
                  21491 1.R1040
5'-most EST
                  LIB3170-051-Q1-K1-B12
Method
                  BLASTX
NCBI GI
                  g4415931
BLAST score
                  401
E value
                  5.0e-39
Match length
                  113
% identity
                  65
NCBI Description
                  (AC006418) unknown protein [Arabidopsis thaliana]
                  >gi 4559393 gb_AAD23053.1_AC006526 18 (AC006526) unknown
                  protein [Arabidopsis thaliana]
```

1975



5'-most EST LIB3051-045-Q1-K1-D2 BLASTX Method NCBI GI g4128208 BLAST score 220 7.0e-18 E value Match length 91 % identity 52 NCBI Description (AF076277) ethylene response factor 1 [Arabidopsis thaliana] >gi 4128210 (AF076278) ethylene response factor 1 [Arabidopsis thaliana] Seq. No. 11432 Contig ID 21494 1.R1040 5'-most EST kl1701213301.h1 Method BLASTX NCBI GI q2809251 BLAST score 1185 E value 1.0e-130 Match length 351 % identity 61 NCBI Description (AC002560) F21B7.20 [Arabidopsis thaliana] Seq. No. 11433 21495 1.R1040 Contig ID 5'-most EST jC-gmro02910040a03a1 Method BLASTX NCBI GI g2462762 BLAST score 762 5.0e-81 E value Match length 248 % identity 62 NCBI Description (AC002292) Highly similar to auxin-induced protein (aldo/keto reductase family) [Arabidopsis thaliana] 11434 Seq. No. Contig ID 21496 1.R1040 sat701012342.hl 5'-most EST Method BLASTX NCBI GI q3193297 BLAST score 761 5.0e-81 E value Match length 184 % identity 71 (AF069298) similar to epoxide hydrolases [Arabidopsis NCBI Description thaliana] 11435 Seq. No. 21497 1.R1040 Contig ID 5'-most EST uC-gmrominsoy057h04b1 Seq. No. 11436 Contig ID 21499 1.R1040

5'-most EST LIB3051-082-Q1-K1-G3

Seq. No. 11437

Contig ID 21500_1.R1040 5'-most EST seb700652885.h1

Contig ID 5'-most EST



```
Method
                  BLASTX
NCBI GI
                  g2827699
BLAST score
                  167
E value
                  1.0e-11
                  57
Match length
                  53
% identity
                  (AL021684) predicted protein [Arabidopsis thaliana]
NCBI Description
                  11438
Seq. No.
                  21500 2.R1040
Contig ID
5'-most EST
                  kl1701212039.h1
Method
                  BLASTX
NCBI GI
                  q2827699
BLAST score
                  167
                  8.0e-12
E value
                  57
Match length
% identity
NCBI Description
                  (AL021684) predicted protein [Arabidopsis thaliana]
Seq. No.
                  11439
                   21501_1.R1040
Contig ID
5'-most EST
                   fC-gmse700752437a2
Method
                  BLASTX
NCBI GI
                  g3309084
                   469
BLAST score
                   8.0e-47
E value
Match length
                   114
                   82
% identity
                   (AF076252) calcineurin B-like protein 2 [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   11440
                   21501 3.R1040
Contig ID
5'-most EST
                   pcp700994873.h1
Method
                   BLASTX
NCBI GI
                   q3309084
BLAST score
                   195
E value
                   5.0e-15
Match length
                   55
% identity
                   67
                   (AF076252) calcineurin B-like protein 2 [Arabidopsis
NCBI Description
                   thaliana]
Seq. No.
                   11441
                   21502 1.R1040
Contig ID
5'-most EST
                   uC-gmropic090g04b1
Method
                   BLASTX
NCBI GI
                   q3287270
BLAST score
                   1044
E value
                   1.0e-114
                   274
Match length
% identity
                   73
                   (Y09533) involved in starch metabalism [Solanum tuberosum]
NCBI Description
                   11442
```

1977

21504 1.R1040

ncj700975676.h1



```
Method
                   BLASTX
                   q3738306
NCBI GI
BLAST score
                   160
                   2.0e-10
E value
Match length
                   76
                   42
% identity
                   (AC005309) unknown protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   11443
                                                                ٠.,:
                   21505 1.R1040
Contig ID
5'-most EST
                   LIB30\overline{5}1-105-Q1-K1-G3
                   11444
Seq. No.
Contig ID
                   21505 2.R1040
                   taw700660105.h1
5'-most EST
                   11445
Seq. No.
Contig ID
                   21507 1.R1040
5'-most EST
                   LIB3051-020-Q1-E1-B12
Method
                   BLASTX
```

NCBI GI g2465925 BLAST score 254

8.0e-22 E value Match length 116 % identity 48

(AF024649) receptor-like serine/threonine kinase NCBI Description

[Arabidopsis thaliana]

11446 Seq. No.

Contig ID 21508 1.R1040

5'-most EST LIB3051-107-Q1-K1-G11

Method BLASTX NCBI GI g3298536 BLAST score 505 E value 3.0e-51 Match length 139 % identity 68

(AC004681) unknown protein [Arabidopsis thaliana] NCBI Description

Seq. No. 11447

21509 1.R1040 Contig ID

LIB3051-070-Q1-K1-F1 5'-most EST

BLASTN Method NCBI GI g3241925 BLAST score 40 4.0e-13 E value Match length 168 % identity 89

NCBI Description Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:

MOK9, complete sequence [Arabidopsis thaliana]

Seq. No. 11448

Contig ID 21510 1.R1040 5'-most EST pxt700943472.h1

Method BLASTN NCBI GI g2264312 BLAST score 37

5'-most EST



```
1.0e-10
E value
Match length
                  322
% identity
                  83
                  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:
NCBI Description
                  MOK16, complete sequence [Arabidopsis thaliana]
                  11449
Seq. No.
                  21510 2.R1040
Contig ID
                  LIB3170-061-Q1-J1-C9
5'-most EST
Seq. No.
                  11450
Contig ID
                  21511 1.R1040
5'-most EST
                  LIB3051-111-Q1-K1-A2
                  BLASTX
Method
NCBI GI
                  g3201627
BLAST score
                  639
E value
                  5.0e-67
Match length
                  139
% identity
                  (ACO04669) putative SWH1 protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  11451
Contig ID
                  21512 1.R1040
5'-most EST
                  LIB3051-112-Q1-K1-A3
Method
                  BLASTX
NCBI GI
                   g2980788
BLAST score
                   434
E value
                   2.0e-42
Match length
                   278
                   43
% identity
                  (AL022197) putative protein [Arabidopsis thaliana]
NCBI Description
                   11452
Seq. No.
                   21514 1.R1040
Contig ID
                   k11701209925.h1
5'-most EST
                   BLASTX
Method
                   g598073
NCBI GI
                   205
BLAST score
                   4.0e-16
E value
                   86
Match length
% identity
                   60
                   (L36806) GT-1 [Arabidopsis thaliana]
NCBI Description
                   11453
Seq. No.
                   21514 2.R1040
Contig ID
                   awf700836550.h1
5'-most EST
Method
                   BLASTX
                   g598073
NCBI GI
BLAST score
                   224
E value
                   3.0e-18
Match length
                   91
% identity
                   62
                  (L36806) GT-1 [Arabidopsis thaliana]
NCBI Description
                   11454
Seq. No.
                   21515 1.R1040
Contig ID
```

1979

jC-gmro02910062b12a1



Method BLASTX
NCBI GI g2088651
BLAST score 295
E value 3.0e-26
Match length 158
% identity 42

NCBI Description (AF002109) hypersensitivity-related gene 201 isolog

[Arabidopsis thaliana]

Seq. No. 11455

Contig ID 21518_1.R1040

5'-most EST LIB3170-041-Q1-J1-F10

Method BLASTX
NCBI GI g4218141
BLAST score 240
E value 2.0e-19
Match length 331
% identity 30

NCBI Description (AJ236702) HMR1 protein [Antirrhinum majus]

Seq. No. 11456

Contig ID 21518_2.R1040 5'-most EST kl1701204524.h2

Method BLASTX
NCBI GI g555655
BLAST score 273
E value 1.0e-23
Match length 159
% identity 55

NCBI Description (U06712) DNA-binding protein [Nicotiana tabacum]

Seq. No. 11457

Contig ID 21518_3.R1040

5'-most EST jC-gmf102220051a02d1

Seq. No. 11458

Contig ID 21518_4.R1040 5'-most EST ssr700556645.h1

Seq. No. 11459

Contig ID 21520_1.R1040 5'-most EST asn701135640.h1

Method BLASTX
NCBI GI 94512667
BLAST score 939
E value 1.0e-107
Match length 311
% identity 64

NCBI Description (AC006931) putative MAP kinase [Arabidopsis thaliana]

Seq. No. 11460

Contig ID 21520_2.R1040 5'-most EST ckk700605794.h2

Method BLASTX
NCBI GI g4512667
BLAST score 171
E value 5.0e-12



97 Match length 56 % identity (AC006931) putative MAP kinase [Arabidopsis thaliana] NCBI Description Seq. No. 11461 21521 1.R1040 Contig ID LIB3170-057-Q1-K1-E2 5'-most EST BLASTX Method g3107903 NCBI GI 186 BLAST score 1.0e-13 E value Match length 61 59 % identity (D83719) polycomb-like protein [Daucus carota] NCBI Description Seq. No. 11462 21522 1.R1040 Contig ID LIB3109-001-Q1-K2-A11 5'-most EST BLASTX Method q3201618 NCBI GI BLAST score 798 2.0e-85 E value 166 Match length % identity 85 (AC004669) Sop2p-like protein [Arabidopsis thaliana] NCBI Description Seq. No. 11463 21522 2.R1040 Contig ID wvk700686490.h1 5'-most EST BLASTX Method q3201618 NCBI GI BLAST score 252 1.0e-21 E value 118 Match length 58 % identity (AC004669) Sop2p-like protein [Arabidopsis thaliana] NCBI Description 11464 Seq. No. 21523 1.R1040 Contig ID LIB3051-027-Q1-K1-A6 5'-most EST BLASTX Method q1850546 NCBI GI 562 BLAST score 9.0e-58 E value 141 Match length 80 % identity (U88045) syntaxin related protein AtVam3p [Arabidopsis NCBI Description thaliana]

11465 Seq. No.

Contig ID 21524 1.R1040

uC-gmflminsoy018e09b1 5'-most EST

BLASTX Method q4539418 NCBI GI 477 BLAST score E value 7.0e-48 153 Match length



```
% identity
                   (AL049171) caffeoyl-CoA O-methyltransferase-like protein
NCBI Description
                   [Arabidopsis thaliana]
Seq. No.
                  11466
                   21525 1.R1040
Contig ID
                  LIB3051-026-Q1-K1-G11
5'-most EST
                  BLASTX
Method
                   q4006848
NCBI GI
BLAST score
                   665
E value
                   6.0e-70
                   168
Match length
                   79
% identity
                   (AJ131433) selenocysteine methyltransferase [Astragalus
NCBI Description
                   bisulcatus]
Seq. No.
                   11467
Contig ID
                   21526 2.R1040
                   rca700999839.hl
5'-most EST
                   BLASTX
Method
                   q3319342
NCBI GI
BLAST score
                   537
                   2.0e-54
E value
                   252
Match length
% identity
                   45
                   (AF077407) similar to mitochondrial carrier proteins (Pfam:
NCBI Description
                   mit_carr.hmm, score: 79.74 and 42.50) [Arabidopsis
                   thaliana]
Seq. No.
                   11468
Contig ID
                   21528 1.R1040
                   jC-gm\overline{l}e01810035d07a2
5'-most EST
                   BLASTN
Method
                   g1370155
NCBI GI
                   120
BLAST score
                   8.0e-61
E value
                   204
Match length
                   90
% identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB11H
                   11469
Seq. No.
                   21528 2.R1040
Contig ID
                   LIB3051-114-Q1-K1-F1
5'-most EST
                   BLASTN
Method
                   g1370155
NCBI GI
BLAST score
                   201
                   1.0e-109
E value
                   333
Match length
 % identity
NCBI Description L.japonicus mRNA for small GTP-binding protein, RAB11H
 Seq. No.
                   11470
                   21529 1.R1040
 Contig ID
```

LIB3093-042-Q1-K1-A10 5'-most EST

BLASTX Method g3395938 NCBI GI 623 BLAST score



6.0e-65 E value 167 Match length % identity 37 (AF076924) polypyrimidine tract-binding protein homolog NCBI Description [Arabidopsis thaliana] 11471 Seq. No. 21529 2.R1040 Contig ID LIB3138-094-Q1-N1-A6 5'-most EST Method BLASTX g3395938 NCBI GI BLAST score 781 2.0e-83 E value Match length 203 % identity

(AF076924) polypyrimidine tract-binding protein homolog NCBI Description [Arabidopsis thaliana]

Seq. No. 11472 Contig ID 21531 1.R1040 5'-most EST LIB3051-032-Q1-K1-G2 Method BLASTX NCBI GI g688423 BLAST score 292 2.0e-26

E value Match length 102 % identity 50

(D26453) tumor-related protein [Nicotiana glauca X NCBI Description

Nicotiana langsdorffii]

Seq. No. 11473

Contig ID 21533 1.R1040

uC-gmrominsoy188e03b1 5'-most EST

Method BLASTX g3461820 NCBI GI BLAST score 286 2.0e-25 E value Match length 62 77 % identity

(AC004138) unknown protein [Arabidopsis thaliana] NCBI Description

11474 Seq. No.

21533 2.R1040 Contig ID epx701106943.h1 5'-most EST

BLASTX Method NCBI GI g3461820 211 BLAST score 1.0e-16 E value 49 Match length 73 % identity

(AC004138) unknown protein [Arabidopsis thaliana] NCBI Description

11475 Seq. No.

Contig ID 21534 1.R1040

jC-gmle01810068f02a1 5'-most EST

Method BLASTX NCBI GI g3522943

BLAST score

Match length

E value

292

197

8.0e-26



```
1300
BLAST score
E value
                  1.0e-144
Match length
                   441
                   36
% identity
NCBI Description
                  (AC004411) putative p-glycoprotein [Arabidopsis thaliana]
                  11476
Seq. No.
Contig ID
                  21536 1.R1040
5'-most EST
                  LIB3109-043-Q1-K1-F1
Seq. No.
                  11477
Contig ID
                  21537 1.R1040
5'-most EST
                  LIB3051-108-Q1-K1-A10
Method
                  BLASTX
NCBI GI
                  g2498565
BLAST score
                  245
E value
                  8.0e-21
Match length
                  111
% identity
                   41
                  C-MYC BINDING PROTEIN MM-1 >gi 1731809 dbj BAA14006
NCBI Description
                   (D89667) c-myc binding protein [Homo sapiens]
                  >gi_4505743_ref_NP_002615.1_pPFDN5_ prefoldin
Seq. No.
                  11478
                  21537 2.R1040
Contig ID
5'-most EST
                  pcp700993248.h1
Method
                  BLASTX
NCBI GI
                   q2498565
BLAST score
                  232
E value
                   3.0e-19
Match length
                  111
                   39
% identity
NCBI Description
                  C-MYC BINDING PROTEIN MM-1 >qi 1731809 dbj BAA14006
                   (D89667) c-myc binding protein [Homo sapiens]
                   >gi 4505743 ref NP 002615.1 pPFDN5 prefoldin
Seq. No.
                  11479
Contig ID
                  21538 1.R1040
5'-most EST
                  LIB3051-110-Q1-K1-G7
Method
                  BLASTX
NCBI GI
                   g4063751
BLAST score
                   395
E value
                   4.0e-38
Match length
                  204
% identity
                   43
                   (AC005851) putative white protein [Arabidopsis thaliana]
NCBI Description
                   >gi 4510409 gb AAD21495.1 (AC006929) putative white
                  protein [Arabidopsis thaliana]
Seq. No.
                  11480
                  21539 1.R1040
Contig ID
                  asn70\overline{1}138037.h1
5'-most EST
Method
                  BLASTX
NCBI GI
                   g3894387
```

Contig ID 5'-most EST

Method

NCBI GI

```
% identity
NCBI Description (AF053995) Hcr2-OB [Lycopersicon esculentum]
                  11481
Seq. No.
Contig ID
                  21540 1.R1040
5'-most EST
                  LIB3167-004-P1-K1-H7
                  BLASTN
Method
NCBI GI
                  a510931
BLAST score
                  292
                  1.0e-163
E value
Match length
                  448
                  92
% identity
NCBI Description V.faba mRNA for alpha 1,4-glucan phosphorylase type H
Seq. No.
                  11482
Contig ID
                  21541 1.R1040
5'-most EST
                  crh700855575.h1
Method
                  BLASTN
NCBI GI
                  q18662
BLAST score
                  535
                  0.0e + 00
E value
Match length
                  579
% identity
                  98
NCBI Description Glycine max hsp 70 gene
                  11483
Seq. No.
Contig ID
                  21542 1.R1040
5'-most EST
                  LIB3109-032-Q1-K1-A5
                  BLASTX
Method
NCBI GI
                  g1172555
BLAST score
                  342
E value
                  4.0e-32
Match length
                  136
% identity
                  47
                  34 KD OUTER MITOCHONDRIAL MEMBRANE PROTEIN PORIN
NCBI Description
                  (VOLTAGE-DEPENDENT ANION-SELECTIVE CHANNEL PROTEIN) (VDAC)
                  (POM 34) >gi 629720 pir S46936 34K porin - potato
                  >gi 1076682 pir A55364 porin (clone pPOM-34) - potato
                  mitochondrion >gi 516166 emb CAA56599 (X80386) 34 kDA
                  porin [Solanum tuberosum]
Seq. No.
                  11484
Contig ID
                  21543 1.R1040
5'-most EST
                  LIB3051-006-Q1-K1-H1
Method
                  BLASTX
NCBI GI
                  g1743354
BLAST score
                  274
                  2.0e-53
E value
Match length
                  201
                  57
% identity
NCBI Description
                  (Y09876) aldehyde dehydrogenase (NAD+) [Nicotiana tabacum]
Seq. No.
                  11485
```

21544 1.R1040

BLASTX

g4539301

dpv701099148.h1



BLAST score 751 E value 1.0e-79 Match length 297 % identity 54

NCBI Description (AL049480) putative mitochondrial protein [Arabidopsis

thaliana]

Seq. No. 11486

Contig ID 21544_2.R1040

5'-most EST uC-gmflminsoy011c06b1

Method BLASTX
NCBI GI g4539301
BLAST score 121
E value 4.0e-11
Match length 97
% identity 48

NCBI Description (AL049480) putative mitochondrial protein [Arabidopsis

thaliana]

Seq. No. 11487

Contig ID 21545 1.R1040

5'-most EST uC-gmflminsoy008g01b1

Method BLASTX
NCBI GI g2723391
BLAST score 240
E value 8.0e-20
Match length 197
% identity 36

NCBI Description (AB007042) EXTR1 [Homo sapiens] >gi_2897905 (AF001690) EXT

like protein 3 [Homo sapiens] >gi_3043562 dbj_BAA25445_ (AB011091) KIAA0519 protein [Homo sapiens] >gi 4103884

(AF029231) EXT homolog [Homo sapiens]

>gi 4503617 ref NP 001431.1 pEXTL3 exostoses

(multiple) -like

Seq. No. 11488

Contig ID 21546_1.R1040

5'-most EST LIB3170-059-Q1-K1-E5

Method BLASTX
NCBI GI g2213594
BLAST score 868
E value 2.0e-93
Match length 257
% identity 64

NCBI Description (AC000348) T7N9.14 [Arabidopsis thaliana]

Seq. No. 11489

Contig ID 21572 1.R1040

5'-most EST LIB3093-052-Q1-K1-D2

Method BLASTX
NCBI GI g1354466
BLAST score 419
E value 6.0e-41
Match length 143
% identity 59

NCBI Description (U52910) U1 snRNP 70K truncated protein [Arabidopsis

thaliana] >gi_1354468 (U52909) U1 snRNP 70K truncated



protein [Arabidopsis thaliana]

11490 Seq. No. 21572 2.R1040 Contig ID 5'-most EST trc700563627.h1 BLASTX Method NCBI GI g1354466 BLAST score 179 2.0e-13 E value Match length 48 75 % identity (U52910) U1 snRNP 70K truncated protein [Arabidopsis NCBI Description thaliana] >gi 1354468 (U52909) U1 snRNP 70K truncated protein [Arabidopsis thaliana] 11491 Seq. No. Contig ID 21574 1.R1040

uC-gmrominsoy194e08b1 5'-most EST

Method BLASTX g452593 NCBI GI BLAST score 210 1.0e-16 E value Match length 129 43 % identity

(D21814) ORF [Lilium longiflorum] NCBI Description

Seq. No.

11492

Contig ID 21575 1.R1040 5'-most EST LIB3051-017-Q1-E1-C10

Method BLASTX NCBI GI g3184283 837 BLAST score 6.0e-90 E value

Match length 239 % identity 71

NCBI Description (AC004136) putative TBP-binding protein [Arabidopsis

thaliana]

11493

Seq. No.

Contig ID 21576 1.R1040 5'-most EST asn701136909.h1

Method BLASTX NCBI GI g3063698 BLAST score 333 E value 1.0e-30 Match length 253 % identity 36

NCBI Description (AL022537) putative protein [Arabidopsis thaliana]

Seq. No. 11494

Contig ID 21576 2.R1040

LIB3051-026-Q1-K1-C2 5'-most EST

Method BLASTX q3063698 NCBI GI BLAST score 181 E value 3.0e-13 Match length 79

NCBI GI

BLAST score



```
% identity
NCBI Description (AL022537) putative protein [Arabidopsis thaliana]
                   11495
Seq. No.
                   21577 1.R1040
Contig ID
                   g4396\overline{5}02
5'-most EST
                   BLASTX
Method
                   g320556
NCBI GI
BLAST score
                   507
                   7.0e-51
E value
                   252
Match length
                   40
% identity
                   chitinase (EC 3.2.1.14) precursor, basic - Arabidopsis
NCBI Description
                   thaliana >gi_166666 (M38240) basic chitinase [Arabidopsis
                   thaliana]
                   11496
Seq. No.
                   21577 2.R1040
Contig ID
                   leu701148149.h1
5'-most EST
                   BLASTX
Method
                   g2980793
NCBI GI
                   421
BLAST score
                   6.0e-41
E value
                   164
Match length
                   47
% identity
                   (AL022197) putative protein [Arabidopsis thaliana]
NCBI Description
                   11497
Seq. No.
                   21578 1.R1040
Contig ID
                   kl1701202926.h1
5'-most EST
                   BLASTX
Method
                   g4406784
NCBI GI
                   203
BLAST score
                   6.0e-16
E value
                   95
Match length
                   48
% identity
                   (AC006532) putative oligopeptide transport protein
NCBI Description
                   [Arabidopsis thaliana]
                   11498
Seq. No.
                   21581 1.R1040
Contig ID
                   seb70\overline{0}654218.h1
5'-most EST
Method
                   BLASTX
                   g2829912
NCBI GI
BLAST score
                   486
                   1.0e-64
E value
Match length
                   158
% identity
                   (AC002291) Similar ATP-dependent RNA Helicase [Arabidopsis
NCBI Description
                   thaliana]
                   11499
Seq. No.
                   21582 1.R1040
Contig ID
                   LIB3093-057-Q1-K1-F9
5'-most EST
                   BLASTX
Method
```

g3297816

Match length



```
E value
                  1.0e-106
                  306
Match length
% identity
                  69
                  (AL031032) putative protein [Arabidopsis thaliana]
NCBI Description
                  11500
Seq. No.
                  21582 3.R1040
Contig ID
                  uC-gmronoir006h07b1
5'-most EST
                  BLASTX
Method
NCBI GI
                  q3297816
                  508
BLAST score
                  2.0e-51
E value
                  149
Match length
% identity
                  64
NCBI Description (AL031032) putative protein [Arabidopsis thaliana]
Seq. No.
                  11501
                  21586 1.R1040
Contig ID
                  jC-gmle01810012e02a1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g1549249
BLAST score
                  315
E value
                  5.0e-30
Match length
                  176
% identity
                  43
                  (U66620) SWI/SNF complex 60 KDa subunit [Mus musculus]
NCBI Description
Seq. No.
                  11502
                  21587 1.R1040
Contig ID
                  ek1700968247.h1
5'-most EST
Method
                  BLASTX
                  g1431629
NCBI GI
BLAST score
                  1396
                  1.0e-155
E value
                  398
Match length
% identity
                   63
                  (X99348) pectinacetylesterase precursor [Vigna radiata]
NCBI Description
                  11503
Seq. No.
Contig ID
                  21587 2.R1040
5'-most EST
                  kl1701211669.hl
Method
                  BLASTN
                  g1431628
NCBI GI
BLAST score
                   35
                  3.0e-10
E value
Match length
                  43
% identity
                   95
NCBI Description V.radiata mRNA for pectinacetylesterase
                   11504
Seq. No.
Contig ID
                   21589 1.R1040
                  LIB3093-023-Q1-K1-F8
5'-most EST
Method
                  BLASTX
                   g2160144
NCBI GI
BLAST score
                   267
E value
                   5.0e-23
```



```
% identity
                  (AC000375) Strong similarity to Arabidopsis oligopeptide
NCBI Description
                  transporter (gb X77503). [Arabidopsis thaliana]
Seq. No.
                  11505
                  21591 1.R1040
Contig ID
                  k11701208328.h1
5'-most EST
Method
                  BLASTX
                  g4262226
NCBI GI
BLAST score
                  347
                  1.0e-32
E value
Match length
                  90
% identity
                  71
                  (AC006200) putative protein kinase [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  11506
                  21593 1.R1040
Contig ID
5'-most EST
                  fua701037592.h1
Method
                  BLASTX
NCBI GI
                  q4006917
BLAST score
                  640
E value
                  1.0e-66
Match length
                  204
% identity
                   58
                  (Z99708) putative protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                  11507
                  21609 1.R1040
Contig ID
5'-most EST
                  crh700850349.hl
Method
                  BLASTN
                   g1113940
NCBI GI
BLAST score
                   233
                   1.0e-128
E value
                   391
Match length
                   90
% identity
NCBI Description
                  Phaseolus vulgaris Pv42p mRNA, complete cds
                   11508
Seq. No.
                   21612 1.R1040
Contig ID
                  hrw701058417.h1
5'-most EST
                  BLASTX
Method
                   g3540199
NCBI GI
                   1196
BLAST score
E value
                   1.0e-132
                   336
Match length
                   68
% identity
                   (AC004260) Putative monosaccharide transport protein
NCBI Description
                   [Arabidopsis thaliana]
                   11509
Seq. No.
```

Contig ID 21613_1.R1040

5'-most EST LIB3051-009-Q1-E1-C10

Method BLASTN
NCBI GI g3941321
BLAST score 246
E value 1.0e-136
Match length 427



```
. % identity
                    Medicago truncatula gamma-glutamylcysteine synthetase mRNA,
  NCBI Description
                    nuclear gene encoding putative plastid protein, complete
  Seq. No.
                    11510
  Contig ID
                    21614 1.R1040
                    epx701110007.hl
  5'-most EST
                    BLASTX
  Method
  NCBI GI
                    q1946359
  BLAST score
                    999
                    1.0e-113
  E value
                    433
  Match length
  % identity
  NCBI Description (U93215) unknown protein [Arabidopsis thaliana]
  Seq. No.
                    11511
  Contig ID
                    21615 1.R1040
  5'-most EST
                    LIB3051-004-Q1-E1-A12
                    BLASTX
  Method
  NCBI GI
                    g2330739
  BLAST score
                    203
                    2.0e-15
  E value
                    120
  Match length
                    42
  % identity
                    (Z98598) putative transcriptional regulator
  NCBI Description
                     [Schizosaccharomyces pombe]
  Seq. No.
                    11512
  Contig ID
                    21616 1.R1040
  5'-most EST
                    jC-qmf102220126b02a1
  Method
                    BLASTX
  NCBI GI
                    g3983665
  BLAST score
                    1551
                    1.0e-173
  E value
                    365
  Match length
  % identity
                    82
                    (AB011271) importin-beta2 [Oryza sativa]
  NCBI Description
                    11513
  Seq. No.
                    21617 1.R1040
  Contig ID
  5'-most EST
                    LIB3051-006-Q1-E1-A8
  Method
                    BLASTX
                    g2739168
  NCBI GI
  BLAST score
                    162
                     4.0e-11
  E value
  Match length
                    102
                     51
  % identity
                     (AF032386) aldose-1-epimerase-like protein [Nicotiana
  NCBI Description
                     tabacum]
  Seq. No.
                     11514
```

Contig ID

21618 1.R1040

5'-most EST

uC-gmrominsoy199c07b1

Method NCBI GI BLASTX g4314401 1345

BLAST score

5'-most EST

Method

```
1.0e-149
E value
Match length
                   342
% identity
                   70
                   (AC006232) putative beta-1,3-glucanase [Arabidopsis
NCBI Description
                   thaliana]
                   11515
Seq. No.
                   21621 1.R1040
Contig ID
5'-most EST
                   uaw700660737.h1
Method
                   BLASTX
                   g2464914
NCBI GI
BLAST score
                   343
                   3.0e-32
E value
Match length
                   101
% identity
                   67
                   (Z99708) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   11516
                   21624 1.R1040
Contig ID
                   LIB3052-009-Q1-N1-D6
5'-most EST
Method
                   BLASTX
NCBI GI
                   q82051
BLAST score
                   108
E value
                   8.0e-11
                   132
Match length
% identity
                   34
                  lipid body-associated membrane protein - carrot
NCBI Description
                   >gi 259453 bbs 117620 (S47635) lipid body membrane
                   protein=DC 59 [Daucus carota=carrots, var Juwarot, Peptide,
                   180 aa] [Daucus carota]
Seq. No.
                   11517
                   21624 2.R1040
Contig ID
                   crh70\overline{0}852419.h1
5'-most EST
                   BLASTN
Method
                   g3097320
NCBI GI
BLAST score
                   38
                   6.0e-12
E value
Match length
                   58
% identity
                   91
NCBI Description Glycine max gene for Bd 30K, complete cds
                   11518
Seq. No.
                   21626 1.R1040
Contig ID
5'-most EST
                   uC-gmrominsoy062c01b1
                   BLASTX
Method
                   g2911052
NCBI GI
BLAST score
                   1021
                   1.0e-111
E value
Match length
                   284
% identity
                   70
                   (AL021961) putative protein [Arabidopsis thaliana]
NCBI Description
                   11519
Seq. No.
                   21627 1.R1040
Contig ID
```

1992

LIB3051-042-Q1-K1-D2

BLASTX

Seq. No.

11524

```
g2072393
NCBI GI
BLAST score
                   830
E value
                   3.0e-89
Match length
                   174
                   92
% identity
                   (U29168) similar to human Xeroderma pigmentosum group B DNA
NCBI Description
                   repair protein, Swiss-Prot Accession Number P19447
                   [Arabidopsis thaliana]
                   11520
Seq. No.
                   21628 1.R1040
Contig ID
5'-most EST
                  ncj70\overline{0}988180.h1
Method
                  BLASTX
NCBI GI
                   g3201635
                   413
BLAST score
                   2.0e-40
E value
                   132
Match length
                   59
% identity
                  (AC004669) hypothetical protein [Arabidopsis thaliana]
NCBI Description
Seq. No.
                   11521
                   21629 1.R1040
Contig ID
                   LIB3051-011-Q1-E1-H5
5'-most EST
Method
                   BLASTX
NCBI GI
                   g3540181
                   497
BLAST score
                   2.0e-50
E value
                   135
Match length
% identity
                   74
                   (AC004122) Unknown protein [Arabidopsis thaliana]
NCBI Description
                   11522
Seq. No.
                   21635 1.R1040
Contig ID
                   LIB3051-002-Q1-E1-E10
5'-most EST
                   BLASTX
Method
                   g728906
NCBI GI
BLAST score
                   266
E value
                   3.0e-23
Match length
                   90
                   56
% identity
NCBI Description
                   PROBABLE CALCIUM-TRANSPORTING ATPASE 5
                   >gi 1077722 pir S50669 hypothetical protein YER166w -
                   yeast (Saccharomyces cerevisiae) >gi 603407 (U18922)
                   Yer166wp [Saccharomyces cerevisiae]
Seq. No.
                   11523
                   21637 1.R1040
Contig ID
5'-most EST
                   LIB3051-034-Q1-K1-E2
Method
                   BLASTX
NCBI GI
                   q1532171
BLAST score
                   496
E value
                   3.0e-50
Match length
                   132
                   70
% identity
                   (U63815) AT.I.24-9 gene product [Arabidopsis thaliana]
NCBI Description
```



```
21638 1.R1040
Contig ID
                  LIB3138-024-Q1-N1-D7
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2213591
BLAST score
                  225
                  3.0e-18
E value
                  135
Match length
                  41
% identity
NCBI Description (AC000348) T7N9.11 [Arabidopsis thaliana]
                  11525
Seq. No.
                  21638_2.R1040
Contig ID
                  LIB3106-031-Q1-K1-H1
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2213591
BLAST score
                  149
                  1.0e-09
E value
Match length
                  64
% identity
                  48
NCBI Description (AC000348) T7N9.11 [Arabidopsis thaliana]
Seq. No.
                  11526
                  21639 1.R1040
Contig ID
                  hrw701062732.hl
5'-most EST
Method
                  BLASTX
NCBI GI
                  g2911072
BLAST score
                  213
E value
                  5.0e-17
Match length
                  88
                  55
% identity
NCBI Description (AL021960) putative protein [Arabidopsis thaliana]
Seq. No.
                  11527
                  21640 1.R1040
Contig ID
5'-most EST
                  uC-gmrominsoy167h12b1
Method
                  BLASTX
                  g3738325
NCBI GI
BLAST score
                  779
                  2.0e-91
E value
Match length
                  188
% identity
                  (AC005170) putative CaMB-channel protein [Arabidopsis
NCBI Description
                  thaliana]
                  11528
Seq. No.
Contig ID
                  21641 1.R1040
5'-most EST
                  pxt700943810.hl
Method
                  BLASTX
                  g3451411
NCBI GI
BLAST score
                  1083
```

E value 1.0e-126 Match length 254 86 % identity

(Z98761) seryl-tRNA synthetase [Helianthus annuus] NCBI Description

Seq. No. 11529

21642_1.R1040 Contig ID



5'-most EST bth700846162.h1

Method BLASTX
NCBI GI g2576361
BLAST score 517
E value 2.0e-52
Match length 144
% identity 66

NCBI Description (U39782) lysine and histidine specific transporter

[Arabidopsis thaliana]

Seq. No. 11530

Contig ID 21645_1.R1040 5'-most EST eep700868213.h1

Method BLASTX
NCBI GI g4008006
BLAST score 501
E value 1.0e-50
Match length 152
% identity 64

NCBI Description (AF084034) receptor-like protein kinase [Arabidopsis

thaliana]

Seq. No. 11531

Contig ID 21646_1.R1040 5'-most EST seb700650524.h1

Method BLASTX
NCBI GI g115473
BLAST score 947
E value 1.0e-102
Match length 253
% identity 71

NCBI Description CARBONIC ANHYDRASE, CHLOROPLAST PRECURSOR (CARBONATE

DEHYDRATASE) >gi_170219 (M94135) chloroplast carbonic anhydrase [Nicotiana tabacum] >gi_445610_prf__1909357A

carbonic anhydrase [Nicotiana tabacum]

Seq. No. 11532

Contig ID 21646 2.R1040

5'-most EST uC-gmrominsoy247g12b1

Method BLASTX
NCBI GI g1168740
BLAST score 581
E value 5.0e-60
Match length 140
% identity 77

NCBI Description CARBONIC ANHYDRASE 2 (CARBONATE DEHYDRATASE 2) >gi 882244

(U19740) carbonic anhydrase 2 [Flaveria linearis]

Seq. No. 11533

Contig ID 21646_3.R1040 5'-most EST jsh701065528.h1

Method BLASTX
NCBI GI g1168740
BLAST score 310
E value 6.0e-35
Match length 131
% identity 60

```
NCBI Description CARBONIC ANHYDRASE 2 (CARBONATE DEHYDRATASE 2) >gi 882244
                   (U19740) carbonic anhydrase 2 [Flaveria linearis]
Seq. No.
                  11534
Contig ID
                  21649 1.R1040
5'-most EST
                  LIB3051-064-Q1-K1-D8
Method
                  BLASTX
NCBI GI
                  g4510339
BLAST score
                  525
E value
                  2.0e-57
Match length
                  142
% identity
                  44
NCBI Description
                  (AC006921) putative ABC transporter protein [Arabidopsis
                  thaliana]
Seq. No.
                  11535
Contig ID
                  21651 1.R1040
5'-most EST
                  LIB3051-104-Q1-K1-A8
Method
                  BLASTX
NCBI GI
                  q3482933
BLAST score
                  436
E value
                  5.0e-43
Match length
                  118
% identity
                  67
NCBI Description
                  (AC003970) Similar to cdc2 protein kinases [Arabidopsis
                  thaliana]
Seq. No.
                  11536
Contig ID
                  21652 1.R1040
5'-most EST
                  uxk700672746.h1
Method
                  BLASTX
NCBI GI
                  g461812
BLAST score
                  506
E value
                  4.0e-51
Match length
                  216
% identity
                  47
NCBI Description
                  CYTOCHROME P450 72 (CYPLXXII) (PROBABLE
                  GERANIOL-10-HYDROXYLASE) (GE10H) >gi_167484 (L10081)
                  Cytochrome P-450 protein [Catharanthus roseus]
                  >gi_445604_prf__1909351A cytochrome P450 [Catharanthus
                  roseus]
Seq. No.
                  11537
Contig ID
                  21654 1.R1040
5'-most EST
                  zpv700757258.h1
Method
                  BLASTX
NCBI GI
                  g4406780
BLAST score
                  1759
E value
                  0.0e + 00
Match length
                  498
% identity
                  80
NCBI Description
                  (AC006532) putative multispanning membrane protein
                  [Arabidopsis thaliana]
```

Seq. No. 11538 Contig ID 21654 2.R1040 5'-most EST dpv701098741.h2